



Doing Economics:

*A Guide to Understanding and
Carrying Out Economic Research*

Steven A. Greenlaw
University of Mary Washington

Houghton Mifflin Company Boston New York

This book is dedicated to my friend Bob McConnell, who showed me this was a feasible project, and to my wife Kathy, who made it possible for me to complete it.

Publisher: Charles Hartford
Editor-in-Chief: George T. Hoffman
Senior Sponsoring Editor: Ann West
Editorial Associate: James Dimock
Senior Project Editor: Kerry Falvey
Manufacturing Coordinator: Karen Banks
Marketing Manager: Mike Schenk

Copyright © 2006 by Houghton Mifflin Company. All rights reserved.

No part of this work may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying and recording, or by any information storage or retrieval system without the prior written permission of Houghton Mifflin Company unless such copying is expressly permitted by federal copyright law. Address inquiries to College Permissions, 222 Berkeley Street, Boston, MA 02116-3764.

Printed in the U.S.A.

Library of Congress Control Number: 2003110136

ISBN: 0-618-379835

2 3 4 5 6 7 8 9—QUH—09 08 07 06

Contents

Preface xi

chapter 1

What Is Research? 1

Research Is the Creation of Knowledge 2

How Are Arguments Evaluated? 3

Summary 5

Suggestions for Further Reading 7

Exercises 8

Appendix 1A: The Range of Economic Methodologies 9

chapter 2

Overview of the Research Process in Economics 11

Research in Science and Nonscience Disciplines 11

Steps of the Research Process for Economics 12

Step 1: Developing an Effective Research Question 14

Step 2: Surveying the Literature 19

Step 3: Analyzing the Problem 20

Step 4: Testing Your Analysis 21

Step 5: Interpreting the Results and Drawing Conclusions 22

Step 6: Communicating the Findings of the Research Project 23

Summary 23

Suggestions for Further Reading 25

Exercises 25

Appendix 2A: Writing the Research Proposal 26

Appendix 2B: Questions for Evaluating a Research Proposal 28

chapter

3

Surveying the Literature on a Topic in Economics	29
Why Is a Literature Survey Necessary?	29
Where to Search: Popular versus Scholarly Literature	30
How to Search: Developing an Effective Search Strategy	32
Browsing	33
Keyword Searching	35
A Basic Search Strategy	38
Sample Keyword Search	39
Obtaining the Resources	40
Summary	42
Suggestions for Further Reading	44
Exercises	44
Appendix 3A: Scholarly References and Citation Styles	46

chapter

4

Using Writing As a Tool for Economic Research	49
Writing to Learn	50
Composition As a Creative Process	50
The Structure of an Argument	51
What Does It Mean to Say That a Conclusion Follows from the Evidence?	53
Examining an Argument	54
Three Types of Reasoning: Deductive, Inductive, and Warrant-Based	57
What Makes for a Persuasive Argument?	59
Example: Evaluating an Argument with Unrepresentative Evidence	62
Example: Evaluating an Argument Using Broad and Deep Reasoning	62
Example: Constructing a Simple Argument	63
An Important Caveat	64
Summary	64
Suggestions for Further Reading	65
Exercises	65
Appendix 4A: Logical Fallacies	67

chapter

5

Writing As a Product of Economic Analysis	71
What Is Economic Writing?	71
Writing Steps	72
Writing the First Draft	72
Features of Good Economic Writing	73
Getting the Ideas Down on Paper	76
Giving Credit for Intellectual Property	77
Revising the Paper	79
Is the Thesis Clear?	80
Is the Paper Well Organized?	80
Are the Points Supported with Evidence?	82
Writing Style	82
Strive for Clarity	83
Use the Active Voice	85
Describe Action with a Verb	86
Be Precise and Concise	86
Writing Mechanics	87
Use Complete Sentences	88
Don't Let Sentences Run On	89
Summary	90
Suggestions for Further Reading	90
Exercises	91

chapter

6

Critical Reading or How to Make Sense of Published Research	93
Making Sense of Published Research	94
Understanding Format	94
Evaluating the Argument: Reading Critically	95
Questions to Guide Critical Reading	97
Evaluating Published Research	112
More Questions to Guide Critical Reading	113

Taking Research Notes and Writing Abstracts and Critical Reviews 117

Summary 120

Suggestions for Further Reading 120

Exercises 121

Appendix 6A: Reading and Evaluating a Theoretical Article 122

chapter



Theorizing or Conceptualizing the Research 125

What Does It Mean to Apply Theory to a Research Topic? 126

What Is Theorizing? 126

Narrative Reasoning 128

Mathematical Reasoning 129

Example 1: Optimizing Models 129

Example 2: Ad Hoc Models 131

A Commonly Used Shortcut: Modifying an Existing Model 132

Example 3: Intertemporal Utility Maximization with Money 132

Example 4: Ad Hoc Example 134

What Makes a Good Research Hypothesis? 134

Summary 136

Suggestions for Further Reading 137

Exercises 137

chapter



Locating (and Collecting) Economic Data 139

Data Creation 140

Steps in Data Creation 140

Sample Data 141

The Structure of Economic Data 141

Characteristics of Data Sets 142

Organizations That Collect and Publish Data 144

Census Bureau 144

Bureau of Economic Analysis 144

Bureau of Labor Statistics 145

The Federal Reserve 145

International Agencies 145

Major Primary Data Collections 145

U.S. National Income and Product Accounts 146

U.S. Flow of Funds Account 146

U.S. Balance of Payments Accounts and International Investment
Position of the U.S. 146

U.S. Census of Population and Integrated Public Use
Microdata Series 147

Current Population Survey 147

Current Employment Statistics 147

The Economic Census 148

Annual Survey of Manufactures 148

Current Industrial Reports 148

American Housing Survey 148

Consumer Expenditure Survey 149

National Longitudinal Surveys 149

Panel Study of Income Dynamics 150

Surveys of Consumers 150

Survey of Consumer Finances 151

Major Secondary Data Collections 151

The Economic Report of the President 151

Economagic 151

FRED II (Federal Reserve Economic Data) 152

STAT-USA/State of the Nation 152

Inter-university Consortium for Political
and Social Research (ICPSR) 152

International Financial Statistics 152

World Economic Outlook Database 153

Penn World Tables 153

Joint BIS-IMF-OECD-World Bank Statistics on External Debt 153

Eurostat 153

OECD Main Economic Indicators and National Accounts 153

Summary 154

Suggestions for Further Reading 155

Exercises 155
Appendix 8A: Overview of Data Sources 157

chapter

9

Putting Together Your Data Set 159

Developing a Search Strategy for Finding Your Data 160

Step 1: Before You Search 160

Step 2: As You Search 161

Data Manipulation 164

Level of Variable 164

Change in Variable 165

Real versus Nominal Magnitudes 169

Index Numbers 170

Quantity Indices Versus Real Quantities 171

Price Indices Versus Implicit Price Deflators 172

How Inflation Distorts Nominal Values 172

Rebasing Data Series 172

Data Smoothing 174

Constructing a Data Appendix to Your Research 175*Summary* 176*Suggestions for Further Reading* 177*Exercises* 178*Appendix 9A: What Are Chained-Dollars?* 179*Appendix 9B: Example of a Data Appendix* 180

chapter

10

A First Look at Empirical Testing: Creating a Valid Research Design 183

Key Issues of Research Design 183**How Does One Analyze Data?** 186

Casual Empiricism 187

Random Variation in Human Behavior 189**Statistical Methods** 192**Simple Statistical Hypothesis Testing** 194

An Alternative Approach: The P-Value 196

Confounding Variables 196**Causal Validity** 198*Summary* 199*Suggestions for Further Reading* 201*Exercises* 201*Appendix 10A: Table of Critical t-Statistics* 202

chapter

11

Introduction to Regression Analysis 203

Steps in Regression Analysis 203

Step 1: State the Hypothesis 204

Step 2: Test the Hypothesis; i.e., Estimate the Relationship 206

Step 3: Interpret the Test Results 208

To What Extent Do the Coefficient Estimates Conform to Your Theoretical Predictions? 209

Statistical versus Economic Significance 212

Are the Coefficient Estimates Plausible? 213

How "Good a Fit" Is the Regression Model? 216

Step 4: Check for and, If Necessary, Correct Common Problems of Regression Analysis 217

Step 5: Evaluate the Test Results 224

Summary 225*Suggestions for Further Reading* 227*Exercises* 227*Appendix 11A: Data Transformations for Regression Analysis* 228*Appendix 11B: Estimation Using Qualitative Variables* 229

chapter

12

Communicating the Results of a Research Project 231

Writing the Research Report 232

Introduction 234

The Written Literature Review 236

Writing the Literature Review 236

Theoretical Analysis 239

Empirical Testing of the Analysis 241

Conclusions	245
Other Components of the Paper	246
Presenting Research Orally	250
An Oral Presentation Is Different from a Written One	251
Preparing the Presentation	252
Use of Visual Aids: Handouts, Transparencies, PowerPoint	253
Practicing the Presentation	256
Giving the Presentation	256
The Role of the Discussant	258
Summary	259
Suggestions for Further Reading	260
Exercises	260
Appendix 12A: Rubric for Grading Research Papers	262
Glossary	263
Reference List	273
Index	277

Preface

"To complete the process of intellectual maturation, every student should be required to apply what he or she has learned to an economic problem and, in the process, acquire experience really 'doing economics.' For a particular intellectual encounter to accomplish this goal, it should involve considerable responsibility on the student's part for formulating questions, gathering information, structuring and analyzing information, and drawing and communicating conclusions to others in an oral and/or written form."

—SIEGFRIED ET AL. (1991)

In the last fifteen years many leading experts in economic pedagogy have argued that rather than merely passing courses, economics students should be required to demonstrate proficiency in *doing* economics. Although doing economics (in the form of economic research) is a key element in graduate education, it has not been the norm at the undergraduate level. Over the last decade this has changed as an increasing number of institutions have added an undergraduate research emphasis to their economics programs. This book aims to guide these novice researchers, whether they are taking an undergraduate course in research methodology, completing a senior capstone course, pursuing a senior thesis, or attempting a research project early in graduate studies.

The objective of *Doing Economics* is to make explicit to novice researchers what experienced researchers know implicitly about the research process. The book's principal thesis is that research is fundamentally a process of constructing persuasive arguments supported by theory and empirical evidence. The more technical issues of literature

searching, theory formulation, data collection, and empirical testing are the details by which the process occurs, but the big picture is about building a case around one's interpretation of an interesting issue or problem.

Research is often viewed as one of those competencies students must have learned in an earlier course or can pick up on their own. But experts in the field of research and writing expose this view as a fallacy.

Although many research skills are taught in traditional courses, the broader skills required for research design tend not to be. Traditional courses succeed in teaching the technical skills—for example, those involving analytic reasoning. When asked to apply a specific theory to an issue or problem, or to perform a specific statistical test on a given data set, most novice researchers do quite well. However, traditional courses are less successful in teaching the broader, creative skills required for research. These broader, creative skills include:

- The ability to frame a good research question when given an issue or problem
- The ability to identify and apply an appropriate model to a given research problem
- The ability to develop an appropriate data set
- The ability to select appropriate ways to test a specific hypothesis
- The ability to use the results to create a scholarly argument

More generally, novice researchers fail to grasp that research is not a straightforward, mechanical process, but one that is replete with ambiguities: there is more than one correct way to approach a given research problem, as well as many incorrect ways. While research can never be reduced to a simple cookbook process, it can be taught in a more explicit manner, and this text seeks to make the steps involved in doing research clear to the novice researcher.

Acknowledgments

This book, more than most I suspect, has benefited from the help of many individuals. Research methodology covers such a range of topics that it is difficult for any author to be a master of them all. I would like to thank the reviewers who commented on this project early on and provided many useful notes and suggestions: Stephen DeLoach, Elon University; Adam Grossberg, Trinity College, Hartford, CT; Wm. Stewart Mounts, Jr., Mercer University; Craig M. Newmark, North Carolina State University; Paul Rappoport, Temple University; and Michael Ye, St. Mary's College of Maryland.

The following people read and commented on one or more chapters of the book, and I would like to thank them as well: John Bean, Seattle University; Gardner Campbell, University of Mary Washington; Lisa Daniels, Washington College; Stephen DeLoach, Elon University; Jane Gatewood, Karen Hartman, and Debra Hydorn, University of Mary Washington; Carol Manning; Linda Manning, University of Ottawa; Kenneth Peterson, Furman University; Jerry Petr, University of Nebraska; Anand Rao and Robert Rycroft, University of Mary Washington; Patrice Scanlon; and Bridget Geiman, Teresa Hannah, Kenneth Lynch, Karen Ransom, Patrick Sarsfield, Stacey Shifflett, Shannon Slawter, and Julie Heselden Topoleski.

This project also benefited from unusually able research assistance from Emily Ruesch and Catherine Stewart.

Matt Kapuscinski, Emily Ruesch, and Cristy Tilghman allowed me to use parts of their research as examples in this book.

Special thanks go to four colleagues who tested the manuscript in their classes: Lisa Daniels, Washington College; Stephen DeLoach, Elon University; KimMarie McGoldrick, University of Richmond; and Kenneth Peterson, Furman University.

I would also like to express my gratitude to Ann West, Kerry Falvey, and the editorial staff at Houghton Mifflin, and finally, to my students since 1996, who were the motivation for this book.

All remaining errors are my own.

S.A.G.

What Is Research?

"I hear and I forget; I see and I remember; I do and I understand."

CONFUCIUS

The best way to *learn* economics is not to hear about it, or to read about it, but to *do* it. Doing economics means performing economic research. As Booth et al. (1995, 2) note, "Doing research can help you understand the material you are studying in a way that no other kind of work can match."

Faculty often view research ability, like writing, as a skill that students either must have learned earlier in their education or can pick up on their own. After all, haven't all undergraduates written term papers? In fact, experts in the field of research and writing have exposed this as a fallacy. Bean (1996), a nationally known scholar of critical thinking, points out that most undergraduates do not really understand how to write a discipline-specific research paper.

This book is intended to remedy this problem. It is designed to guide students through the research process from the conception of the research question to the completed research report. Students often find that research is a difficult process. As they struggle to complete the project they conclude that something is wrong with *them*. What they fail to grasp is that research *is* difficult. It is not a straightforward, mechanical process but one replete with ambiguities, wrong turns, and "wasted" efforts. Progress in the research process often comes in fits and starts rather than evenly. It isn't always or even usually predictable. In contrast to most of the assignments students encounter in their education, with research there may not be a well-defined answer. Every researcher, including experts, feels confused or even overwhelmed at some point during a research project. The difference is that experienced researchers, like marathon runners hitting "the

wall," know that this is normal and that with perseverance one will get through it.¹ In short, this book is designed to make explicit to undergraduate economics students what experienced researchers know implicitly about the research process.

The book is an introduction to research methodology in economics. Methodology is more than a fancy word for methods. Research methods are the tools of economic research: for example, online search skills, critical reading and writing, and statistical methods. Research methodology is how those tools are combined to produce valid and reliable research.

Research Is the Creation of Knowledge

One popularly held view of research is that it means "the search for knowledge." This definition seems to view knowledge as like fruit on a tree in a forest. Thus, all that is necessary in the research process is for the researcher to discover the tree and collect the knowledge, much like a farmer picks fruit.

This idea of knowledge is based on the traditional view of science as entirely objective, where the data, when collected, "speak for themselves." In other words, the data will yield the same conclusions to any researcher.

This view fails to differentiate between knowledge and facts. Knowledge is what is believed to be true about something, what is believed to be a correct understanding of something. Facts are just data. Knowledge, by contrast, is facts with meaning; that is, it is an expert's best interpretation of the facts. You can see this distinction if you compare the results from a chemistry lab (i.e., the data) with the researcher's discussion of the results in a lab report (i.e., the interpretation). Only the latter is the knowledge.

Research, then, is not merely searching for facts. Research is more completely defined as the creation of (valid) knowledge.² Facts are *discovered*; knowledge, as an interpretation, is *created*.³ Instead of a farmer harvesting a fruit tree, a better analogy for research is a detective searching for clues and then developing a case on the basis of those clues or other evidence. Scholars create knowledge by constructing arguments. In a research context, an argument is not primarily a quarrel or controversy. Rather, an argument is a point of view or position on a question. More formally, an argument is an assertion or claim supported by reasons or evidence.⁴ Knowledge in any discipline can be thought of as a conversation or dialog between scholars as they develop competing arguments. One example familiar to students of macroeconomics is the conflicting views between the Monetarists and the Keynesians. Through this dialog, arguments are evaluated, the weaker ones are winnowed out, and the stronger ones are refined and improved.⁵ Thus, over time knowledge in the field advances.

Students and even some faculty believe that it is simply unrealistic to expect undergraduates to be able to perform original research. Cohen and Spencer (1993, 222) cite student comments such as "How can I tell you anything you don't already know?" and "How can you expect an undergraduate to say anything original?"

Part of the reasoning behind such comments stems from the perception that research includes only groundbreaking, paradigm-shifting examples like the theory of evolution or the theory of relativity. There are two problems with this perception. First, the majority of research represents only marginal improvements in our understanding. This is not to say that such improvements are not important, merely that they represent relatively modest advances in our knowledge. Second, even major "breakthroughs" are based on the work that came before. For example, Ethridge (1995) points out that the "discoveries" we associate with such notables as Alfred Marshall and John Maynard Keynes can be traced back to earlier work by lesser-known scholars.

In fact, undergraduates *can* complete serious, legitimate research projects. This is attested to by the growing number of undergraduate economics programs that require senior research projects and the increasing number of journals that publish undergraduate research.⁶ Booth et al. (1995, 7) assert that "it is no exaggeration to say, that done well your [research] will change the world tomorrow."

Ethridge (1995) distinguishes between two types of research: discovery, that is, "formulating, finding, and creating new knowledge or information" and confirmation, that is, "discerning the validity or reliability of knowledge or information." Undergraduate research may be largely confirming, such as running new tests of previously established theories. Such research is still considered "new" in that it adds to our knowledge by, for example, applying the previous theory to new data or new situations, which if successful, broadens its applicability.

How Are Arguments Evaluated?

Earlier, we defined research as the creation of *valid* knowledge, and we described knowledge as a dialog between competing arguments. Arguments "compete" through their validity. What makes knowledge valid? Each discipline has its own approach and language of discourse, but they all boil down to the use of logic and evidence to support a conclusion.⁷

Let's explore this idea in more detail, to find out how arguments are evaluated and either (provisionally) accepted or refuted.

First-year college students tend to think "**dualistically**."⁸ That is, they perceive (nearly) every question to have a unique correct answer, like a problem

in mathematics, for example, $2 + 2 = 4$. Students think in terms of black and white, despite the fact that we live in a gray world. This perception is not their fault; they've been trained for at least twelve years to think that way.

By the time students are sophomores and juniors, their thinking has advanced to "multiplicity." That is, questions have more than one correct answer. For example, in macroeconomics students learn both that "The Great Depression was caused by inappropriate monetary policy" and that "The Great Depression was caused by instability in the private sector." Most students conclude that since multiple points of view exist, knowledge and truth are essentially subjective! Though multiplicity is a more sophisticated form of thinking than dualistic thinking, it's not valid to conclude that knowledge and truth are subjective. In truth, the fact that there are different points of view about an issue or question does *not* imply that all points of view are correct or equal. (I have often said to my students, "There is more than one correct answer to this question, but there are also an infinite number of incorrect answers.") Suppose you are a juror in a criminal trial and both the prosecution and defense bring in experts who contradict each other. There are two points of view, but only one is true: the defendant is either guilty or innocent. How do you decide?

To make a valid decision, you will need to think critically about the testimony.⁹ Critical thinking is one of those concepts (like "liberal education") that college students are familiar with but often find difficult to put into words. Missimer (1995) defines critical thinking as the evaluation of competing arguments on the basis of their evidence. Thus, it involves the ability to recognize and assess an argument and its constituent parts, including the assumptions, logic, and evidence. This will be explained in detail in Chapter 4. For now, we can say that when scholars evaluate different arguments, they ask questions such as:

What are the reasons behind the argument?¹⁰

Does the argument make sense? Why or why not?

Is the logic flawed?

What are the underlying assumptions (explicit and implied)? Are they flawed?

How critical are the assumptions? That is, would different assumptions lead to different conclusions?

What is the empirical evidence? Does it support the conclusion?

In light of the reasons and evidence provided, is the argument persuasive? If so, the conclusion is valid.¹¹

Let's think about how this winnowing process works. Arguments are evaluated to see if they stand up to intellectual scrutiny according to the

criteria just given. If flaws are discovered, the arguments are disproved, after which they may be revised or discarded. For the moment, however, they are removed from consideration. If flaws are not discovered, the arguments are only provisionally accepted, never proved.

That arguments can be disproved but never proved is sometimes difficult for people to grasp. Consider a sports metaphor. Who is the best tennis player in the world? Players compete and as long as they win, they remain in the running for the title. When they lose, they are (at least temporarily) out of contention. But no player can ever be the best forever, only provisionally so, until a better player comes along. It is the same way with arguments.

Think of an argument as a rope across the Grand Canyon. Each test of an argument is like a thread. No one test is strong enough to bear your weight. Over time, however, when multiple tests confirm the argument, the thread becomes a string, the string becomes a rope, and the rope becomes a thick cable. At some point, you become willing to trust it. By contrast, imagine that when a test of the argument fails, the thread is cut. If there are enough failures, the rope is weakened, and ultimately severed in two.

In short, we evaluate the supporting evidence so as to weed out flawed arguments. If we are left with a single point of view, we can conclude that it is correct. Often, however, there is more than one argument left.¹² How do we choose then? When scholars face more than one argument they can choose between them on the basis of disciplinary norms, such as efficiency or equity.¹³ For example, consider the following two positions:

Free international trade is better for a nation than protectionism because under free trade a nation will have more goods and services, that is, a higher GDP.

Tariffs and quotas are necessary to protect workers and businesses because certain sectors of the economy cannot compete internationally under free trade.

Both are arguable: the first on the basis of efficiency and the second on the basis of equity.

SUMMARY

- Research is the search for or creation of valid knowledge.
- Rather than a collection of facts, knowledge is interpretation of the facts.
- The community of learners in any discipline represents a dialog between competing arguments: assertions made about the correct interpretation of the facts.

- How do arguments compete? What makes for valid knowledge?
- Valid research is research that is demonstrable and persuasive to other scholars in the field.
- Research is persuasive when it is done systematically, by design, and where the conclusions are supported by reason and empirical evidence.

NOTES

1. Booth et al. (1995, 23–25) make a similar point.
2. Ethridge (1995) defines research as “the systematic approach to obtaining new and reliable knowledge” (p. 16).
3. Truth is not created, but our understanding, our interpretation, is created.
4. Missimer (1995).
5. See Appendix 1A for a more detailed treatment of how knowledge is created in economics.
6. Chapter 2 provides specific examples of outlets for undergraduate research in economics.
7. Ethridge (1995) defines validity as “demonstrable to others based on reason and evidence.” This is a useful way to think of it. Remenyi et al. (1998, 24) argue similarly that a “researcher has to be able to convince an audience of the value and relevance of his or her research efforts. . . . In addition, the academic researcher needs to explain why his or her research should be considered important and needs to be able to point out precisely what was found and what use the findings are to the community.”
8. This view of thinking, including the labels “Dualistic” and “Multiplicity,” is from Nelson’s (1989) interpretation of Perry (1970).
9. Ethridge (1995) notes that critical thinking is an essential element for research.
10. Missimer (1995) calls assertions that are unsupported by reasons “loose arguments.”
11. Nelson (1989) characterizes this level of thinking as “Contextual Relativism.”
12. This is analogous to the concept of Pareto optimality, as demonstrated, for example, by a production possibilities frontier, where all of the

“points” are Pareto efficient but to choose between them requires some additional criteria for judgment.

13. Nelson (1989) labels this level of thinking as “Contextually Appropriate Decisions.”

SUGGESTIONS FOR FURTHER READING

- Bean (1996)*, especially Chapter 12—Useful guide to teaching research, writing, and critical thinking to undergraduates. Chapter 12 is a rethinking of how to use research papers as an effective tool for undergraduate education. Bean emphasizes the view of research as developing an argument.
- Booth et al. (1995)*—Classic text on college-level research. High-level research text: not concerned with mechanics, but rather research design and argument construction. Focus is on research in the humanities but is very applicable to research in the social sciences as well.
- Blaug (1992)*—Well-known book-length survey of the range of methodologies used by economists. For a similar but shorter survey, see Hausman, 1989.
- Ethridge (1995)*—Readable guide to research methodology in economics at the graduate level. Ethridge teaches agricultural economics; as such, the book is very applied and thus comprehensible to undergraduate researchers. See especially Chapters 1–4 on methodological foundations of economic research.
- Friedman (1968)*—Thoughtful article on research methodology as practiced by economists. For a similar article, with a slightly different point of view, see Machlup (1965).
- Hausman (1989)*—Survey article on economic methodology. For a more detailed survey, see Blaug (1992).
- Machlup (1965)*—Another view on research methodology as practiced by economists. Compare with Friedman (1968).
- Nelson (1989)*—Widely used summary of Perry’s taxonomy of critical thinking.
- Perry (1970)*—Classic study of the levels of critical thinking obtained by U.S. undergraduates.
- Remenyi et al. (1998)*—Guide to research methodology in business administration. Designed for master’s- or doctoral-level research, but as an applied text, it has much to offer undergraduate researchers.

EXERCISES

1. Economists have a reputation for being unable to reach a consensus on issues. You may have heard the joke that if you laid all the economists in the world end to end, they would fail to reach a conclusion. This reputation is due, in part, to the way economic knowledge is created. The reference list for this chapter includes articles by Fritz Machlup and Milton Friedman, two respected economists who address this issue. Read the articles by Machlup (1965) and Friedman (1968). According to each author, what are the major reasons for disagreement among economists? What can you infer from this list of reasons about how economic knowledge (i.e., agreement) is created?

The Range of Economic Methodologies

We observed that scholars in any discipline present competing arguments to resolve issues. This is also true of those who study economic methodology. The view presented in this chapter represents a consensus of the majority of economists, but the dialog of economic methodologists includes a range of views.

Hausman (1989) argues that real-world complexities in the social sciences, such as researchers' inability to conduct controlled experiments, preclude any chance of convincing empirical evidence. Thus, scholars should use inductive reasoning to establish the basic psychological or technological laws of economics, such as utility maximization or diminishing marginal productivity.

Then, for a specific application of such a law, one can deduce the economic implications. For example, assuming consumers maximize utility, one can deduce that higher prices result in fewer items purchased.

At the other end of the methodological spectrum is Blaug (1992), who argues for falsification. That is, not only should hypotheses be judged on the basis of empirical testing, but scientists should actively try to refute the hypotheses of their research. They should reject those that are refuted, instead of refining and retesting ad nauseam.

The view presented in this chapter is closer to Blaug than to Hausman.

One view, which is somewhat outside this spectrum, is McCloskey (1998). She argues that economists don't rely merely or even primarily on scientific forms of proof such as logic and empirical evidence, but on literary ones as well, such as introspection, case studies, and metaphor. That economists use these devices is clear; whether or not they can persuade other experts of their views without using scientific forms of proof is not.

Overview of the Research Process in Economics

"If we knew what we were doing, it wouldn't be called research, would it?"

ALBERT EINSTEIN

In the previous chapter, we defined research in fairly generic terms. While we used examples from economics, the story could have been applied to almost any discipline, from anthropology to zoology. Now let's focus more closely on the research process in economics. This chapter will provide an overview of that process, with an emphasis on the first step: developing an effective research question. Subsequent chapters will examine the remaining steps of the process in detail. At the end of the introduction to each step in this chapter, you will be referred to the appropriate chapter later in the book.

At this stage, most of the material introduced in this chapter will seem like Greek to you. Though you may have already been exposed to some of what we discuss here, until you complete your own economics research project you are unlikely to really grasp it. That's okay! Everyone in your class is in your shoes. And every professional researcher has gone through this process. You can be sure that by the end of your project, everything will make a great deal more sense.

Research in Science and Nonscience Disciplines

Academic knowledge can be divided into two classes: science disciplines and nonscience disciplines. The difference between science and other disciplines lies both in the types of questions they explore and the kinds of proof they accept as convincing. Scientific questions are fundamentally questions of objective fact: what causes earthquakes or how do catalysts work? Nonscientific questions are questions about the affective realm:

what is the meaning of life or how does a sunset make you feel? Science is research that is empirically testable, at least in principle, while research in the nonscientific disciplines is not empirically testable. This is not to say that there are no standards of proof in nonscientific disciplines, merely that the proof is normative and not based on empirical facts.

Scholars in the sciences perform research using the scientific method. The **scientific method** is a set of procedures for drawing valid, reliable, and objective conclusions. These procedures include the following steps:

- Select a scientific problem or question;
- Apply a theory to derive a hypothesis about the problem or question;
- Test the hypothesis by comparing its predictions to evidence from the real world;
- If the hypothesis fails the test, modify it (and retest) or reject it;
- If the hypothesis passes the test, provisionally accept it;
- Test the hypothesis in a new context.

The scientific method is not an arbitrary set of procedures. Rather, for research to be valid it needs to be done systematically. Spector (1981, 7) observes, “It is not necessary to use scientific methods to answer [research] questions. One might rely on intuition or educated opinion.”¹ Isn’t this how most people answer questions, especially those for which they lack specialized knowledge? The problem with such nonscientific methods is that they are subject to one’s biases and other subjective factors. Psychologists have identified the “fallacy of personal validation,” in which people tend to “see” what they hope or expect to see in evidence. The scientific method, while not infallible, is designed to avoid or at least minimize these biases, and result in objective findings.

Steps of the Research Process for Economics

Researchers in economics, as a social science, use a version of the scientific method. This research process in economics can be broken down into six steps, shown in Table 2.1.

As we discuss these six steps, you will see that a research project is not a linear process; rather, the steps are overlapping and iterative. Instead of completing the steps one by one until the end, often the researcher finds it necessary to return to an earlier step in the process as he or she revises his or her thinking.² For example, though you need to begin thinking about the research question early in the process (Step 1), after you have reviewed what other scholars have done in the area (Step 2) you may need to revise the question. Similarly, in determining how to perform the empirical testing (Step 4), you may need to consult previous testing method-

NOTES FOR NOVICE RESEARCHERS

Requirements for a Good Research Project

There are a number of things necessary for completing a good piece of research. These include:

- A Good Research Question (Chapter 2)
- A Testable Hypothesis (Chapter 7)
- A Good Data Set (Chapters 8–9)
- An Empirical Methodology that Adequately Tests the Hypothesis (Chapters 10–11)

(Each of these is discussed in detail in the chapters indicated in parentheses.)

If your project lacks any of these four items, the best you can hope for is a mediocre outcome in your research. It makes sense then, at least early on in your project, to think about these items in parallel, rather than sequentially, since any one of them can trip you up. For example, as soon as you identify a research question, you should begin looking for data in that area, even if that means reading ahead in the appropriate chapters of this book.

Note also that these four things are necessary, but not sufficient, for a good project. In other words, they don’t guarantee a good project will result. In addition, you will need to correctly interpret your empirical results as well as persuade the reader of your conclusions.

Table 2.1 The Research Process in Economics

1. Developing an Effective Research Question
2. Surveying the Literature on the Topic
3. Analyzing the Issue or Problem
4. Testing Your Analysis
5. Interpreting the Results and Drawing Conclusions
6. Communicating the Findings of the Research Project

ologies (Step 2). Additionally, lack of adequate data for testing (Step 4) may require you to reconceptualize the problem (Step 3) in a way that is testable. Finally, it always makes sense to begin writing the research report (Step 6) as you progress through the research process, while the various steps are fresh in your mind, rather than waiting until the end. In short, it may be more accurate to describe research as a looping process. Indeed, Ethridge (1995) observes, "Research is a creative process. . . . There is no 'magic formula' for accomplishing or producing good economic research. If it were that simple, anyone could do it." In this respect, performing research cannot be reduced to following a recipe. Nonetheless, it is helpful to think about the process in terms of discrete steps.

Step 1: Developing an Effective Research Question

At the start of the process, the researcher must define the scope of the project. This requires answering, at least tentatively, three questions:

1. What is the research topic?
2. What is the research question?
3. What is the research hypothesis?

Let's define these terms. The **research topic** is the general area the project will cover. For example, the topic might be "unemployment in the United States." The **research question** is the specific focus of the research. For example, "How do different levels of education affect unemployment rates?" The **research hypothesis** is the researcher's proposed answer to the research question or the researcher's principal assertion about the topic, which will be supported by the paper.³ For example, "Individuals with higher levels of education will tend to have lower levels of unemployment." You may have to start with a tentative hypothesis until you get further into the research. That is normal. The procedure for determining a research hypothesis is explained below in Step 3.

Notice that, grammatically speaking, a research topic is a subject, while a research question must be a complete sentence. The same is true of the research hypothesis, which must also be a complete sentence. This is not merely grammar—a subject may indicate only a vague, general idea of the research, while a complete sentence shows a greater depth of thinking about it.

During the planning phase, the most important of these three concepts is the research question. A good one will make the research project much easier and more likely to be successful. A poor one will, at best, make the project more difficult. A good research question has several characteristics, listed in Table 2.2. Let's elaborate on each of these.

Table 2.2 Characteristics of a Good Economic Research Question

- Problem-oriented
- Analytical (rather than descriptive)
- Interesting and significant
- Amenable to economic analysis
- Feasible, given the time and resources available

A research question is almost always better when it is framed in terms of a problem or concern to be addressed. This tends to focus the question in a useful way, which is less likely to happen otherwise.⁴ Closely related to the problem orientation is a second characteristic: a good research question should be analytical rather than factual. That is, its purpose should be to explain some aspect of the problem rather than describe it. For example, instead of asking a question like "What is poverty?" one should ask something like, "Why is the poverty rate among children so much higher than the average poverty rate among Americans?" Bean (1996) calls descriptive papers "all about" papers. In other words, they discuss "all about" a topic without necessarily having a clear point or argument. Booth et al. (1995) observe that factual questions tend to ask who, what, when, or where, while analytical questions tend to ask why or how?

Readers tend to react to descriptive papers by asking "So what?" In other words, they don't find the papers interesting or significant. Let's turn to these two concepts next.

A research question should be interesting both to the researcher and to the audience. It should be interesting to the researcher because if it is not, he or she will find it difficult to become engaged by the problem.⁵ Booth et al. (1995, 36) state, "Nothing will contribute to the quality of your work more than your sense of its worth and your commitment to it." Your interest will help you stay motivated until the project's completion, especially if it is a long project.

More importantly, a good research question should be interesting to one's audience. To grasp what this means in the context of research, you need to understand the concept of a research community. Booth et al. (1995, 17–18) provide a very helpful definition: A research community is the collection of scholars who perform research on a topic. These are the experts in the field who teach courses as well as publish books and papers

NOTES FOR NOVICE RESEARCHERS

Can Undergraduates Perform Real Research?

Your instructor knows that you are not a professional researcher and so he or she won't hold you to that high a standard. At the same time, undergraduates can and do perform real, original research. In short, they can contribute to the conversation of scholars on a topic, so you shouldn't sell yourself short or set too low a bar for yourself.

on the subject.⁶ These scholars may be at your university or at other institutions anywhere around the world. They know one another by virtue of hearing each other's research presentations and by reading each other's published work. A research question should be of interest to this audience because if it is not, the research project is not worth doing. Remember, the point of research is to contribute to the conversation of scholars on an issue.

At this point, student researchers face a real problem. They actually have two audiences: the theoretical audience just described and a practical one consisting of their classmates, their instructor, and anyone who might read their paper if it is submitted for publication. We will make some suggestions about how to deal with this problem later, but for now, suffice it to say that a good research question is one that interests your readers.

When we talk about scholars being "interested" we mean more than that they think the question is entertaining. In a research context, interest is closely related to the concept of **significance**. Is the question a "hot topic," an area of current interest to the research community? Booth et al. (1995, 18–19) note, "If your question is already a live issue in your community, most readers will care about it when you pose it. . . . [Indeed, they] will be especially interested if you can convince them that they don't understand [it] as well as they thought they did." However, if your topic is not a "live issue," you will have to actively convince your audience that the question is worthy of study.

Booth et al. (1995) define significance by asking how weighty is the problem on which the research focuses. The greater the weight of the problem, that is, the higher the costs of leaving it unresolved, then the more significant the problem is from a research perspective. Consider three examples. Poverty among able-bodied individuals who choose not to work is not considered to be a very significant problem. Poverty among the elderly who

NOTES FOR NOVICE RESEARCHERS

Example of a Good Research Question: To What Extent Does Parental Employment Adversely Affect Children's Cognitive Development?

This is an example of a good research question. To understand why, consider how the criteria from Table 2.2 apply:

- **Problem-oriented**—In the last generation, the labor force participation rate among women has increased dramatically. As a result, the typical U.S. family is no longer characterized by one parent (usually the father) working, while another parent (usually the mother) stays home with primary responsibility for child rearing. Rather, it is increasingly the case that both parents are employed, and the children are placed in child-care. Given that more and more children are receiving child-care by someone other than a parent, does this adversely impact children's cognitive development? If so, perhaps families will need to revisit the issue of employment by both parents.
- **Analytical**—To address this research question, one would need to study the process of children's cognitive development and, in particular, the role that parents or other child-care providers play. This requires an analytical approach.
- **Interesting and Significant**—Past studies have indicated the possibility that children's cognitive development is adversely affected by the lack of parental input. However, these studies have not developed a consensus on either the magnitude or permanence of this effect. Clearly, researchers have more to learn about this question, which affects many American families.
- **Amenable to Economic Analysis**—Noneconomists may wonder about the relevance of economic analysis to children's cognitive development, a topic that is more normally associated with the field of educational psychology. However, this topic is a straightforward application of labor economics and human capital theory. In particular, economists studying this issue have applied the theory of an educational production function, by which parental care and other inputs such as spending on food, shelter, health care, and the like are combined to "produce" educated children.

- **Feasible**, given the time and resources available—The issue of feasibility obviously depends on the researcher and the context in which the research is being done. Suffice it to say here that there are a number of excellent data sets that can be used in this research, including the National Longitudinal Study of Youth and the Panel Study of Income Dynamics–Child Development Survey.

were not able to save adequately for retirement could be considered a significant problem. Poverty among children, for which they are not responsible and which may continue into their later lives, is probably a very significant problem. The ultimate test of significance, say Booth et al. (1995), is the extent to which a research paper asks readers to change their beliefs.

What makes an **economic** research question? Recall from your first-principles course that economics was defined as a content area (i.e., involving questions that originate from the general condition of scarcity), but more importantly as a way of thinking. Is the question or problem that the research will examine amenable to economic analysis? In other words, can economic analysis give us some insights into the problem? For example, does the question deal with choice under constraints? Does it involve demand or supply, needs, wants, availability? If the answer to any of these questions is yes, then you have an economic research question.

Finally, a researcher should be able to complete investigation of the research question in a reasonable period of time, for example, a semester or an academic year, given the resources available. A primary consideration here is the availability of adequate data to assess the hypothesis. Some researchers are fortunate enough to begin their project with a rich data set. For them, the challenge is determining an interesting and significant research question they can explore with that data. Most researchers, however, start with the question and then need to search out suitable data. For further information on obtaining data, see Chapters 8 and 9.

Students doing a research paper for a class often find it a bit difficult to select a research question—and for good reason: the assignment is artificial. In the real world, a business economist is told, “Here is a problem—research it!” Similarly, an academic economist does not say, “Gee, I have to do a research project, so let me think of a topic.” Rather, as an expert in some field(s), he or she is aware of where the research frontier is, what is

known, and what issues or problems in the field have yet to be satisfactorily explained. Those are the research questions he or she is likely to choose. This suggests a strategy for selecting student research questions:

1. Pick a general topic area that interests you, ideally one in which you have some background (e.g., an area in which you have taken a course).
2. Start reading the literature, not merely to see what has been done, but also to identify what questions remain to be answered or what problems remain to be solved.
3. Select a promising research question from what you have found in the literature. For example, could an interesting previous study be applied to a new place or time? Alternatively, are there conflicting findings on some question? If so, you might try to reconcile the conflict by examining the issue again, using a different set of data or a different methodological approach. Studies conclude by suggesting questions for further research. Perhaps one of these questions would be feasible. Finally, the literature survey may reveal gaps in the literature that you could explore.

Examples of good research questions that undergraduates have investigated in recent years include:

- What factors determine the demand for blood in the Fredericksburg, Virginia, region?
- Do union jobs systematically pay higher wages than nonunion jobs when disaggregated by occupation?
- Does class attendance influence student performance in macro principles courses?
- How much of a premium are consumers willing to pay for organic vegetables?
- How much are consumers willing to pay for cleaner streams in the Enoree watershed of South Carolina?
- Do nearby polluted streams affect housing prices?

Step 2: Surveying the Literature

The **literature** on a topic refers to the research that has been completed on that topic. When a researcher reviews or surveys the literature, he or she is asking the following questions:

- What is currently known?
- What has been discovered to date on a given topic?

Thus, the objective of a literature survey is for the researcher to identify and become familiar with the major studies that have been published on

a topic. Only after completing the literature survey can he or she finalize the research question. We will discuss surveying the literature in more detail in Chapter 3.

Step 3: Analyzing the Problem

The objective of a research project is to explain some aspect of a significant issue or problem. This step is that explanation. It is the conceptual or theoretical analysis of the issue or problem, where theory is applied so as to shed light on the problem.

Many students find this to be the hardest part of the research process—indeed, some students explicitly omit this part of the project. This is a significant flaw in their research, since the “analysis” is the central part of the process. What does it mean to “apply theory to an issue or problem”? Theorizing sounds abstract and difficult, but this step need not be. By definition, a theory is a simplified version of reality that allows the researchers to more easily analyze the issue or problem. As a researcher, you need to ask yourself: what are the essential concepts comprising the problem being researched? (The nonessential concepts can be ignored.) How are the essential concepts related? What do these relationships imply? The result of this process of analysis or conceptualizing is the research hypothesis.

Probably the single most important determinant of how successful your research will turn out is the hypothesis you work with. Recall that the hypothesis is the proposed answer to the research question. This suggests a way to start.

If the research question is:

What caused U.S. consumer spending during the 1990s to increase at a faster rate than income?

A tentative hypothesis could be:

U.S. consumer spending during the 1990s increased at a faster rate than income because . . .

So how does one fill in the dots? How does one figure out a reasonable hypothesis? One way is to derive the hypothesis from theory, possibly with ideas derived from the literature. What economic theory or theories are applicable to the current problem? There is a well-developed theory of consumer spending. What does this consumption theory suggest affects consumer spending in addition to income? Theory suggests that household wealth and interest rates influence consumer spending. How did those factors change during the 1990s? Interest rates were roughly stable over the last half of the decade; stock markets, however, boomed. Does it appear that one or more factors can explain the change in consumer spending? If so, that could be your hypothesis:

U.S. consumer spending during the 1990s increased at a faster rate than income because of the increase in household wealth caused by booming stock markets.

We will discuss conceptualizing the model in more detail in Chapter 7.

Step 4: Testing Your Analysis

In terms of research methodology, there is a major divide between science and nonscience disciplines. Science is research that is empirically testable, at least in principle, while research in the nonscientific disciplines is not empirically testable. This is not to say that philosophy, literature, or any of the other humanities is less valuable or less valid than biology, chemistry, or other sciences. All this implies is that the humanities use nonempirical methods to validate their scholarship.⁷

Economics as a social science uses a version of the scientific method in research.⁸ What does it mean to test a theory empirically? The researcher asks the question, If the theory is correct, what evidence should there be of it in the real world? For example, if the theory of supply is valid, then if the price of some product rises, producers should supply greater quantities of output. In other words, testing a theory means comparing the implications or predictions of the theory with appropriate real-world evidence. The data are collected and analyzed. If the theory fails, then it is modified and retested, or it is rejected. If the theory does not fail, it is provisionally accepted.

Take, for example, what happened to unemployment during the Great Depression. By 1933, the unemployment rate had increased to 25 percent. What did the existing theory suggest? The microeconomic theory of labor markets predicted that when faced with a decrease in demand, unemployment would be only temporary until real wages fell. Though nominal wages fell, so did the price level. As a result, it is not clear whether real wages fell or not, but unemployment remained above 10 percent until 1942! Since the prediction of the theory (only temporary unemployment until real wages fell) was contradicted by the real-world evidence (significant unemployment for over a decade), the theory was shown to be inadequate.

For this step in a research project, there are really several elements. The first is deciding how you will go about assessing the validity of the hypothesis. This is an important part of designing or planning the research. Research design has two parts. The first is locating a good data set, one that is large, rich, and accurate enough to adequately test your hypothesis. We noted earlier that failure to find a good data set would compromise your ability to complete your research successfully. For this reason, it is important to begin looking for data early in the planning process. A detailed explanation on data sources will be found in Chapters 8 and 9.

The second part of research design is selecting a statistical method to adequately test your hypothesis. A key question is the following: if the statistical test you employ yields the best possible results, how confident can you be that your hypothesis is confirmed? In other words, will the test adequately discriminate between your hypothesis and alternatives? If not, then you should consider a more powerful test. Oftentimes, however, lack of adequate data will limit your ability to use a more powerful test. This underscores again the importance of finding a good data set. A discussion of statistical methods for testing hypotheses will be found in Chapter 10.

Once you have collected your data and settled on an appropriate statistical test, the next element is *doing* that assessment. This typically requires no more than running a computer program.

Step 5: Interpreting the Results and Drawing Conclusions

This step is closely related to the previous one, but it is sufficiently important that it needs to be emphasized separately. Ethridge (1995, 29) argues that this step “is the one most often overlooked or underestimated by students as they are being initiated to the research process. They often see the empirical computer results as the end product of the process. Actually, the interpretation of those results may be more challenging, and more important, than the process of obtaining the numbers.”

What are the results of the empirical testing of the research hypothesis? Are they consistent with the predictions of the theory? Are there any problems in the testing protocol that need to be corrected to obtain valid results? What are the shortcomings of the testing methodology that limit or weaken the results? In light of the answers to these questions, what can be concluded about the results? To what extent do they support the hypothesis? How do the results compare with those of earlier studies? More generally, what can be concluded about your analysis and about the research question more broadly?

To be sure, there are numerous additional details to consider when interpreting statistical results. These details will be discussed in Chapter 11.

We need to speak to one final but often misunderstood point. What is good research? Good research is research that follows the scientific method, wherever the results lead, even if they reject the hypothesis. Recall that research is the search for valid knowledge. Ethridge (1995) observes that honest research is open minded; looking for a preconceived conclusion is not honest research. A research project that rejects the hypothesis is not a failed project because it still advances our knowledge—in this case, by eliminating one hypothesis as an explanation for the problem.⁹

Step 6: Communicating the Findings of the Research Project

The final step in the research process, and in many ways the most important one, is communicating the results. This is how knowledge progresses. The communication begins with the researcher's written report of the research. This is where the researcher makes a case for the validity of his or her hypothesis based on the logic and empirical evidence of the research. At one level, the research paper really does no more than explain what the researcher discovered from each of the steps in the research process. It is important to remember, however, that the research does not stand on its own; in other words, the evidence is not self-evident. Rather, the purpose of the research report is to present a convincing answer to the research question, in a way that persuades other researchers in the field.

Again, like everything else in the research process, the research report is done in several stages. Typically, a draft is written. The researcher asks colleagues to review it and provide suggestions for improvement.

The researcher may present an oral version of the paper at a professional conference, again to receive comments from other experts. For example, economics students have the opportunity to present their research at a number of venues, including the annual meetings of the Virginia Social Science Association, the Eastern Economic Association, and the National Conference on Undergraduate Research. After each of these stages, the researcher's paper is revised to incorporate what has been learned from the feedback received.

Finally, the paper is submitted to a journal for publication. For example, several journals exist to publish undergraduate economics research. These include *Issues in Political Economy* and *University Avenue Undergraduate Journal of Economics*. Usually, the journal editors ask for still more revisions before a researcher's paper is published.

SUMMARY

The research process in economics has six major steps:

1. Developing an Effective Research Question
2. Surveying the Literature
3. Analyzing the Problem
4. Testing the Analysis
5. Interpreting the Results and Drawing Conclusions
6. Communicating the Findings of the Research Project

NOTES

1. Ethridge (1995) makes the same point when he observes that there are six ways to gain knowledge: through the senses, experience, intuition, revelation, measurement, and reasoning. Of these, only the last two are sources of reliable data.
2. This book is designed to be used this way too. Though laid out sequentially, it includes many references backward and forward.
3. Note that the hypothesis of your research will become the thesis of your research paper.
4. Note that a research problem is not necessarily a social or other problem in the real world. For further elaboration of this point, see Booth et al. (1995), especially Chapter 4. Bean (1996, 30) describes how scholars identify a research problem: "Expert[s] feel an uncertainty, doubt a theory, note a piece of unexplained data, puzzle over an observation, confront a view that seems mistaken, or otherwise articulate a question or problem."
5. On a project such as an honor's or master's thesis or a doctoral dissertation, it is critical to choose a research advisor whose interests are in your topic area; otherwise, they will find it difficult to give you the commitment you need.
6. Members of a research community can be identified by reviewing the published literature.
7. Missimer (1995) describes these methods as "speculation." This term is not meant to be dismissive; it merely means a type of reflection similar to the conceptualizing that occurs in the theoretical sciences, where validity is based on logic and normative standards, but not empirical evidence.
8. We will discuss the difference between the social and physical sciences in Chapter 10 on empirical methods.
9. This is also consistent with Blaug's (1992) doctrine of falsification.
10. Another analogy that might make sense to undergraduates is your plan for completing degree requirements.
11. In a class setting, the audience consists of your classmates and the instructor; in a professional setting, the audience may consist of a funding agency.

SUGGESTIONS FOR FURTHER READING

- Booth et al. (1995), Chapters 2–4*—Excellent discussion of what makes a good research question.
- Ethridge (1995)*—Readable guide to research methodology in economics. See especially Chapters 5–9 on designing and carrying out a research project.
- Ramanathan (1995), Chapter 14*—Good summary of the major steps involved in completing an empirical research project, from choosing a topic to writing the report.
- Siegfried et al. (1991), the section titled "Purpose of the Economics Major," pp. 199–202*—Concise but complete statement of the research methodology of economists, which they label the "economic way of thinking."
- Wyrick (1994)*—Guide to economic research for undergraduates. Published just before the Internet revolution, but much of the book is useful. See, especially, Chapters 9–12.

EXERCISES

1. Consider the following sample research questions. Using the criteria discussed in this chapter for a good research question, determine which are good research questions and which are poor research questions. Explain your reasoning.
 - a. Why is the supply of hospitals low in certain areas, and how does this affect the quality and price of the hospitals and doctors in these areas?
 - b. Why do consumers buy over-the-counter dietary supplements that are said to promote weight loss despite the health risks, and what factors influence these purchases?
 - c. What were the causes of the Great Depression?
 - d. What is the effect of foreign direct investment on a nation's economic growth rate?
2. Design a good research question. Explain how it meets the criteria for a good research question.
3. Write a complete research proposal, following the model described in Appendix 2A. You will likely have to read ahead in this book to complete a good proposal.

Writing the Research Proposal

A successful research project requires planning. This is particularly true for inexperienced researchers, who often find planning the project difficult. This difficulty is all the more reason to spend time and effort on it.

Your research plan should be flexible. You can't expect the project to go exactly according to the plan, but without a plan it's hard to know where you will end up. Think of it as the blueprints for a house you are building.¹⁰ It is not that you can't or won't make changes as you go along, but you don't want to start the project without a plan. To develop a plan, a scholar must think about the research process.

The research plan becomes tangible in the form of the research proposal. A proposal is an exercise in persuasion, whereby you seek to convince your audience that you can complete a meaningful and interesting piece of research on the topic you have selected.¹¹ In the proposal, you sketch out in detail *how* you propose to do that. Thus, the proposal describes the research plan. Ethridge (1995, 85 and 89) points out that "The more thorough and complete the proposal, the clearer and more complete the plan" and also that "The proposal is evidence of the quality of thought that has gone into the project, and the whole research plan . . . will be judged by the quality of the project proposal."

A research proposal should have five components:

1. **Statement of the Nature of the Problem:** What is the issue or problem this research will address? Why is it interesting, significant, and amenable to economic analysis?
2. **The Research Question:** What is the question that is the focus of the research?
3. **Survey of the Literature:** Outline the major studies in the literature that have a bearing on the research question. Explain clearly how they are relevant to your research question. (For a more extensive discussion of the written literature survey, see Chapter 12.)
4. **Research Design:** What is the analytical framework of the model? For example, "This research will apply the theory of demand to study consumer spending on Hula-Hoops." Note that the research hypothesis should follow from this conceptual analysis.

In an ideal world, what data would you need to test your hypothesis? For example, sales of Hula-Hoops, their price, personal disposable income, some measure of their appeal to consumers.

What data have you found to test your analysis, and what is the source of that data? For example, "I plan to test the hypothesis using data on Hula-Hoop sales and prices, as well as disposable income obtained from the National Income and Product Accounts, Survey of Personal Consumption Expenditures. I was unable to find a measure of coolness."

How will you test your analysis empirically? For example, "This research will conduct a regression study over the time period of 1972–1999." Explain how this design should produce reliable and valid results.

(Detailed explanation of these topics will be found in Chapters 7, 8, 9, 10, and 11, respectively.)

5. **References:** List all the references *you have read so far and plan to use in your paper*. Do not list references you have not read or references that provide background information but do not have a clear bearing on your research question. Use the appropriate bibliographic style.

In sum, the proposal should answer four questions. What is the issue or problem your research will investigate? What specific question will you attempt to answer? How will you attempt to answer it? Do you have adequate sources to do the job? If the proposal does not convince the reader that it answers all four questions sufficiently, it is not an acceptable proposal.

If you have a good research proposal, completing the research project will be straightforward—you just need to follow the plan. Even if the plan changes, the project should be doable with a minimum of problems. If you have an incomplete or otherwise poor proposal, it is a signal that the project will be difficult or impossible to accomplish successfully.

Questions for Evaluating a Research Proposal

We observed above that a key purpose of a research proposal is to convince the reader that the researcher can complete a solid piece of research on an important topic. When you evaluate a research proposal you should consider the following questions and issues:

- Does the proposal explain the problem or issue sufficiently to make it interesting and worthy of a reader's time? Why or why not? Is there anything in the introductory section of the proposal that ought to be explained in more detail?
- Do you understand the research question? In other words, can you explain what the research will attempt to do or prove? Is there a clear statement of this? If so, identify it. If not, say so.
- Does the author appear to have identified the major studies previously done on the topic? Has he or she identified those studies' contributions and shortcomings? (The proposed research should build on the contributions and/or correct one or more of the shortcomings.)
- Do you understand *clearly and exactly* what the author intends to do to complete this research? Are there any details missing from the description? If so, what are they? Has the author identified a reasonable data set with which to test his or her hypothesis? Has he or she adequately explained his or her testing methodology?
- Are there enough sources to write a convincing paper?
- In sum, what do you see as the proposal's strengths, and what do you see as its weaknesses?

Your commentary should conclude with an explicit statement about whether the proposal is acceptable as is or not. If it is not, you should provide specific guidelines for what needs to be done to make the proposal acceptable. Again, remember that your feedback should be supportive but honest.

Surveying the Literature on a Topic in Economics

*"If I have seen further, it is by standing
on the shoulders of Giants."*

SIR ISAAC NEWTON

Earlier we noted that scholars engage in a kind of conversation or dialog of competing arguments as they attempt to increase humankind's knowledge about a topic. This conversation manifests itself as published research, which is described as "the literature" on the subject. When you survey the literature on a subject, you are trying to identify the major studies that have been published to date. But more importantly, you are trying to understand what the studies say and how they relate to one another. We will explain how to read and understand scholarly publications in Chapter 6. In this chapter we will explain why and how researchers perform a literature survey as a key component of the research process. In Chapter 12, we will explain how to write a literature survey as part of a research paper.

Why Is a Literature Survey Necessary?

Ethridge (1995, 115) observed that before you can "advance the state of knowledge," you need to know what the state of knowledge is. So when you survey the literature, you are trying to create your own sense of what is known and what is not known about the subject. This is important for several reasons. Early in the research process you can get ideas for possible angles for your research. Later in the process you can show how your research fits into and contributes to the larger conversation on the subject. This serves to establish your credentials as a researcher by showing readers that you are knowledgeable about the field.

There are several additional reasons for surveying the literature. One is to avoid duplicating findings that have already been obtained. In this fashion you can use someone else's prior work as a steppingstone for yours. For example, Ethridge (1995) points out that the literature can help you design your own study by showing how previous approaches either were or were not successful. Similarly, previous studies can suggest ways you might handle problems you run into.

Those new to research sometimes focus their literature survey too narrowly. The literature review should not be limited to searching for previous studies on your *topic*. It should also examine studies of other topics that use conceptual frameworks or testing methodologies that might be applicable to your topic. For example, in a research project on the demand for blood, it could be useful to examine previous demand studies on related products, such as the demand for transplant organs or even on other commodities entirely.

Where to Search: Popular versus Scholarly Literature

Information on a topic can be found in two types of publications: popular and scholarly publications. Scholarly publications are also referred to as scientific or professional publications. What is the difference between popular and scholarly publications? Popular publications are addressed to a general audience, while scholarly ones are addressed to a specialized audience, namely, experts in the field. Scholarly publications are often primary sources of information, while popular publications are almost always secondary sources. A primary source of information is the original publication of a research study, written by an expert and addressed to other scholars in the field. One example would be an article in the *American Economic Review* (AER, a scholarly journal) that explains a new study conducted by the authors of the article on the effects of preschool education on poverty rates. A secondary source is a report based on the primary source, typically addressed to a more general audience. An article in *Time* magazine summarizing the article in the *American Economic Review* would be an example of a secondary source in a popular publication.

One popular publication that undergraduates sometimes mistake for a scholarly journal is the *Economist*. This magazine publishes a regular column entitled "Economics Focus" that does an excellent job of summarizing recent research in language that the educated lay public can understand. In addition, unlike most popular publications, the *Economist* often provides references to the scholarly articles in which the research it reports on was published. Nonetheless, the *Economist* is a popular rather than a scholarly journal.

To begin the literature survey process, it makes sense to consult popular sources because they will give you information on what is generally known about a topic. In addition, since they are addressed to general audiences they are easy to read and assimilate. Popular sources include encyclopedias; news magazines, such as *Time*, *Newsweek*, and *Business Week*; and newspapers such as the *Wall Street Journal* or the *New York Times*. Note that these popular publications are increasingly available online, often for free. For example, the most recent two weeks of the *Washington Post* are available at www.washingtonpost.com.

Internet resources such as directories and search engines are also popular sources of general information. Most undergraduates are familiar with these tools, so we will only discuss them briefly here. Directories are hierarchically organized catalogs on a variety of subjects. Well-known examples include <http://www.galaxy.com>, <http://www.looksmart.com>, and <http://www.yahoo.com>. Virtual libraries are directories that have been compiled by librarians and other experts. As a result, they tend to include somewhat better resources than ordinary directories. Examples include the WWW Virtual Library (<http://www.vlib.org>) and the Internet Public Library (<http://www.ipl.org>). Each of these has sections on economics that can provide helpful background information.

Another Internet tool is the search engine. Search engines are computer programs that create databases of websites catalogued by topic. Users can perform keyword searches on those databases to obtain information from the websites dealing with their subjects of interest. Popular search engines include <http://www.google.com>, <http://www.excite.com>, and <http://www.altavista.com>. A particularly useful search engine is MetaCrawler at <http://www.go2net.com>. Instead of searching its own database, MetaCrawler searches the results of several major search engines, including Google, Yahoo, Ask Jeeves, About, LookSmart, and others.

The shortcoming of most general-interest directories and search engines is that they contain a great deal of information that is not useful for scholarly research. This is especially problematic for novice researchers who may have difficulty differentiating between authoritative and nonauthoritative sources. We will discuss some ways in which these tools may be productively used later in this chapter.

Although it makes sense to begin reviewing the literature by consulting popular sources, it is important to realize that the conversation between experts in the field that creates knowledge occurs only in the scholarly literature. These professional publications have undergone a peer referee or review process prior to publication. This review conveys at least a certain degree of validity to these studies. Scientific publications include articles

Table 3.1 Examples of Scholarly Economics Journals

National Journals:	<i>American Economic Review</i> <i>Journal of Political Economy</i> <i>Quarterly Journal of Economics</i> <i>Econometrica</i> <i>Review of Economics and Statistics</i> <i>Review of Economic Studies</i>
Regional Journals:	<i>Economic Inquiry</i> <i>Southern Economic Journal</i> <i>Eastern Economic Journal</i>
Specialized Journals:	<i>Journal of Urban Economics</i> <i>Journal of Money, Credit and Banking</i> <i>Journal of Industrial Economics</i>

in professional journals, scholarly books, monographs, working papers, and some government documents. Monographs are short books, typically no more than one hundred pages, on a single (hence “mono”) scholarly topic. Working papers are scholarly articles in progress—prior to final publication. Table 3.1 gives examples of scholarly economics journals published in the United States.¹

Many professional journals are difficult for undergraduates to understand, and as we noted earlier, Chapter 6 will explain how to decipher them. You should be aware that some journals are more readable than others.² For example, each regional Federal Reserve Bank publishes an *Economic Review*. Publications of the Congressional Budget Office are also quite good. In addition, publications of the various think tanks such as the Brookings Institution, the American Enterprise Institute, the Economic Policy Institute, and the Cato Institute tend to be more accessible. However, it is important to remember that think tanks are in business to promote their particular point of view. Thus, one should read their publications more carefully and critically than more objective sources.

How to Search: Developing an Effective Search Strategy

To locate information efficiently, one needs to use a search strategy. This section will explain how to develop one by using a sample research question: “To what extent was the 2001 U.S. economic slowdown caused by the decline in the stock market?”

Table 3.2 AEA/JEL/EconLit Subject Descriptors

A – General Economics and Teaching
B – Schools of Economic Thought and Methodology
C – Mathematical and Quantitative Methods
D – Microeconomics
E – Macroeconomics and Monetary Economics
F – International Economics
G – Financial Economics
H – Public Economics
I – Health, Education, and Welfare
J – Labor and Demographic Economics
K – Law and Economics
L – Industrial Organization
M – Business Administration and Business
N – Economic History
O – Economic Development, Technological Change, and Growth
P – Economic Systems
Q – Agricultural and Natural Resource Economics
R – Urban, Rural, and Regional Economics
Z – Other Special Topics

There are two general approaches to locating information. Ackermann and Hartman (1998) describe them as browsing and keyword searches. You can think of them as emphasizing human capital (or thinking) versus physical capital (or brute force). Neither is superior for all searches; rather a key task in developing a search strategy is to determine the optimal combination of the two for a given search.

Browsing

Browsing means manually examining a document (e.g., a printed bibliography or an online directory) for useful information or references to useful information. Effective browsing requires you to think carefully and logically about how information about a discipline is organized. The American Economic Association (AEA) uses a hierarchical system to classify subject information in economics, an overview of which is shown in Table 3.2.³

Table 3.3 Expanded AEA Subject Descriptors
for A – General Economics and Teaching

A00 – General

A1 – General Economics

- A10 – General
- A11 – Role of Economics; Role of Economists
- A12 – Relation of Economics to Other Disciplines
- A13 – Relation of Economics to Social Values
- A14 – Sociology of Economics
- A19 – Other

A2 – Teaching of Economics

- A20 – General
- A21 – Pre-college
- A22 – Undergraduate
- A23 – Graduate
- A29 – Other

Each of these general headings can be further expanded, as illustrated in Table 3.3 for the category “A – General Economics and Teaching.” The complete classification system can be reviewed at <http://www.aeaweb.org/journal/elclasjn.html>. By knowing this classification system, you may be able to find useful information on a research topic more easily.⁴

There are two journals you should be particularly aware of as you begin your research. The *Journal of Economic Perspectives* (JEP) is designed as a general-interest journal for economists. As such, the majority of articles it publishes are summaries of current thinking by economists on a given topic. Similarly, the *Journal of Economic Literature* (JEL) primarily publishes survey articles of the literature on specific topics in economics. The American Economic Association publishes both journals, and their tables of contents for the past several years are available at the AEA website (<http://www.aeaweb.org/jep/contents> and <http://www.aeaweb.org/journal/contents>).⁵ If you can find an article on your research topic in either of these journals, you will be ahead of the game since they will almost certainly identify and discuss the relevant major studies as of the publication date of the article. You will still need to look for more recent studies, however.

More generally, when you locate a useful journal article, you should consider the sources it references as obvious targets for your review. Bear

in mind, however, that since the article’s author may not have pursued the same research question as you, its references should be a complement to, *not* a substitute for, your own literature survey. In sum, browsing is more likely to give you useful information than will simple keyword searches, but your results will depend on the judgment of the author or editor of the document you browsed.

Sample Browse Search

Suppose you are researching the extent to which the recent U.S. economic slowdown was caused by the decline in the stock market. You are aware that the life-cycle model of consumption theory suggests that household wealth is an important determinant of consumer spending. A review of the contents of the *Journal of Economic Perspectives* reveals an article by James Poterba titled “Stock Market Wealth and Consumption” in the spring 2000 issue. The article includes two pages of references going back to Ando and Modigliani’s classic 1963 paper on which the life-cycle model is based.

Keyword Searching

The alternative way to locate information is by performing a keyword search. Keyword searches use search engines on the World Wide Web or on specialized databases. These latter include bibliographic databases, such as *EconLit* (formerly the *Economics Literature Index*), which consist of citations and abstracts only, as well as full-text databases, such as Dow Jones Interactive, which contain entire documents. Access to specialized databases is provided by a number of commercial vendors, such as OCLC FirstSearch or DIALOG. Researchers can obtain access to many of these databases through their university library or computer network.

Keyword searching allows you to examine far more documents than would be feasible by browsing, even assuming you knew about them. The major web search engines claim to index all the documents on the Internet, while search engines in electronic databases index all the documents in the database.⁶ For example, *EconLit* currently indexes more than six hundred economics journals, seventeen hundred new books, and nine hundred new dissertations per year—for nearly a complete record of all English-language publications in economics since 1969. As such, *EconLit* should be one of the first places you start searching when researching in economics.

One exceptional tool for keyword searching is the *Social Science Citation Index* (SSCI). The SSCI allows you to search for citations both backward and forward in time. In our earlier discussion of browsing, we noted that when you find a useful publication, say DeLoach (2001), it makes sense to review

NOTES FOR NOVICE RESEARCHERS

Do Not Limit Your Search to Full-Text Databases!

Undergraduates sometimes view the literature search as nothing more than a hurdle on the way to completing a research project. (This is not true—indeed, as we have tried to explain, discovering what other researchers have done in the field can be a major help in completing your own research.) As a consequence, such students may limit their search to full-text databases. After all, they reason, even if you find something in a bibliographic database, it's still possible it would be difficult or impossible to acquire. This is poor judgment! The goal of a literature search is to find the most important studies done on a topic, not merely the easiest ones to obtain. Put another way, if there is a major study that you fail to find because it was not cataloged on a full-text database, your research will have a substantial flaw.⁷ Nearly all college and university libraries subscribe to the major economics journals. This means articles from those journals are available in printed form, microfiche, or film at your library. Even if your library does not subscribe to a journal that you need, it's possible that one of your faculty members does. It never hurts to ask. Limiting yourself to full-text databases will also virtually rule out any books as sources. Finally, it is hard to imagine a reference that you could not obtain through interlibrary loan. In short, don't limit yourself to full-text databases! It's false economy—it will cost you in the end more than it's worth.

the earlier studies that it cites. *SSCI* can do this for you electronically, since for every journal article or book it catalogs, it also indexes the publication's citations. Doing this electronically can be a bit faster than reading a study's bibliography, but the real advantage of *SSCI* is in doing forward searches. That is, the *SSCI* database also indexes all the publications that cite DeLoach (2001) in *their* reference lists. As a result, it is very easy to get a complete list of the major studies done on a topic.

An additional advantage of the *SSCI* is that it is interdisciplinary; it covers a variety of social science fields. What this means is that you are less likely to miss a study done on your topic but in a field other than economics, as could occur if you only searched *EconLit*. The main disadvan-

tage of the *SSCI* is that it is quite expensive, so many colleges and universities do not subscribe to it. If your institution has it, you should use it.

Though your scholarly literature survey should concentrate on using specialized databases to find literature, web search engines can also be profitably employed. The best web search engine for academic uses is probably Google (<http://www.google.com>). A criticism of Web-based research is that the majority of what is on the Web is not scholarly in nature. One advantage of Google is that in addition to indexing webpages, it also catalogs a number of other file types, including Adobe document (.pdf) files, Word and WordPerfect documents, Excel and Lotus 123 spreadsheets, PowerPoint presentations, and others. As a result, a researcher could use Google to find scholarly working papers or data sets on researchers' personal webpages, even before they are published.

The principal disadvantage of keyword searches, when using specialized directories or web search engines, is that they may generate many "hits," but only a few useful ones. Imagine doing an electronic search using the phrase *stock market*. Depending on the database, you could get hundreds or even thousands of results. Even if there are many relevant hits, you will need a great deal of time and effort to separate the wheat from the chaff. Consider the way commercial fishermen use a dragnet. The net scoops up everything in the water for several miles around. Then even though the fishermen only want codfish, they need to examine and throw away everything else they catch.

The solution to this quandary is to perform a more advanced keyword search using a combination of Boolean and phrase searching. Boolean searching allows you to focus the search in the way most likely to obtain useful results. There are three Boolean operators: AND, OR, and NOT.

AND is used to **narrow a search**. For example, a search on the keywords *Keynesian* AND *Post* would locate only those items including **both** keywords, such as "Post Keynesian."

OR is used to **widen a search**. A search on the keywords *Monetarist* OR *Keynesian* would locate those items including **either** keyword.

NOT means **exclude**. It includes all items **except** those that contain the NOT keyword. For example, a search on the keywords *Keynesian* NOT *Post* would identify all items with "Keynesian," except those with "Post Keynesian."

You can also fine-tune your search using nested Boolean logic. For example, a search on the keywords *Monetarist* OR (*Keynesian* NOT *Post*) would locate all items including either "Monetarist" or "Keynesian," except those with "Post Keynesian."

Phrase searching looks for the items that contain an exact phrase that you specify within quotes. For example, a search on “*Post Keynesian*” would locate only those items with that exact phrase, but it would exclude items that include those two keywords separately. A wildcard character can allow you to search more efficiently by truncating your search phrase. If the wildcard character for a search engine is *, then searching on the keyword *Keynes** would locate items containing “Keynes” or “Keynesian” or “Keynesians.” Similarly, a search on the keyword *wom*n* would yield items containing “woman” or “women.”

A Basic Search Strategy

Ackerman and Hartman (1998) propose a basic search strategy, which is shown (in a revised form) in Table 3.4. Though the table gives the impression that this is a straightforward process, in practice it is a great deal more uncertain. For one thing, at the beginning of the process the researcher may not know exactly what to look for. This is especially true for a researcher new to the topic. For another, the process is iterative. You will never find everything in a single search! Rather, you will need to try an initial set of keywords with one database. Once you find a few good sources, read them. As you become more familiar with the literature, other ideas for keywords will come

Table 3.4 A Basic Search Strategy

1. Begin by stating your research topic or question;
2. Identify important concepts relevant to your topic;
3. Brainstorm to create a list of keywords that describe these concepts;
4. Determine any synonyms to these keywords;
5. Choose a subject approach; that is, determine which discipline or disciplines are likely to have literature on your topic; for example, economics, law, public policy, business, sociology, education;
6. Determine which search features may apply; for example, Boolean operators, wildcards;
7. Choose an appropriate database to search the given subject; for example, *EconLit*;
8. Read the search instructions for the database;
9. Create a search expression using the appropriate syntax;
10. View the results;
11. Modify the search if necessary (return to Step 2);
12. Try the same search with another database (return to Step 6).

to mind, which you can then search with. If one database is unsuccessful, try another. If one subject approach doesn't yield enough, try a related field. There is an art to keyword searching, and some researchers are better than others at it. Fortunately, anyone can become better with practice.

Sample Keyword Search

Let's perform a keyword search on our sample research topic using the search strategy outlined in Table 3.4.

1. Recall that the research topic is **the extent to which the 2001 economic slowdown was caused by a decline in the stock market**.
2. The research topic can suggest the important concepts to begin with. These could include **stock market decline** and **economic slowdown**.
3. A little brainstorming about possible connections between the two can suggest other keywords to search for: **household wealth**, **consumer spending**, and **life-cycle model**. Note also that in choosing keywords, one can select the JEL subject classification for one's topic. In this case, we could choose *G100* for “General Financial Markets” or *E200* for “Macroeconomics: Consumption, Saving, Production, Employment, and Investment, General.”
4. Synonyms for (stock market) *decline* could be *crash* or *correction* or *bear market*. A synonym for (economic) *slowdown* could be *recession*. Synonyms for *wealth* could be *savings* or *saving*. Synonyms for *consumer spending* could be *consumer expenditure* or *consumption*.
5. The research topic suggests two fields whose literature might prove useful: economics and business.
6. For multiple-word expressions we will need to use phrase searching, for example, “*consumer spending*.” For synonyms we will need to use Boolean operators, for example, “*consumer spending*” OR “*consumer expenditure*” OR “*consumption*.” For relationships we will need to use the Boolean AND operator, for example, “*consumption AND wealth*.”
7. When performing economics research, a good database to begin with is *EconLit*. For other fields, the following might be helpful:
 - *Expanded Academic ASAP* – for business academic journals,
 - *Wilson Business Abstracts* – for business academic journals,
 - *Business & Company Resource Center* – for business information at the company level,
 - *ERIC* – for education,
 - *PAIS* – for public affairs and international studies,
 - *Dow Jones Interactive* – for current topics not yet in the scholarly literature, also business academic journals,
 - *SSCI* – for any of the social sciences.

Be aware that the user interface and the format in which search results are provided can differ depending on which vendor supplies your access to the database.

8. The search instructions for *EconLit* are available at <http://www.econlit.org/econlit/hints.html>.⁸
9. An initial keyword search was performed using the expression “*consumption AND wealth*.”
10. This search yielded 1,422 hits, which is too many to review meaningfully.
11. A revised search using the expression “*consumption AND stock market*” yielded 97 hits, including a promising article by Shirvani & Wilbratte (2000), “Does Consumption Respond More Strongly to Stock Market Declines than to Increases?” A subsequent search using the expression “*consumer spending (OR “consumption”) AND recession*” resulted in 78 hits, including another good one, Blanchard (1993), “Consumption and the Recession of 1990–91.” Another search using the expression “*stock market AND recession*” yielded 13 hits of which three look useful.
12. Let’s try searching another database: Expanded Academic ASAP. A search using the expression “*consumer spending (OR “consumption”) AND recession*” yielded 22 hits, of which several appear interesting. One of these was Blanchard (1993), which we found previously.

You should try all the search expressions that generated good results (i.e., not too many hits, and many of them useful) on *each* of the databases searched. At some point, you will begin to turn up the same sources, like Blanchard (1993). This is a sign that you’re near the end of effective searching. Notice that our keyword searches did not turn up the Poterba (2000) paper we discovered earlier by browsing. That is why both types of searches are necessary. After completing the browsing and keyword searches, and examining the references cited by the studies found there, we can conclude that we have found the major published studies to date on our research question.

Obtaining the Resources

Once you have identified potential sources, you need to obtain them. The World Wide Web has revolutionized access to research materials, especially for scholars at smaller institutions whose libraries may have only limited collections. More and more resources are available online, including full-text books and journal articles.

More than a hundred economics journals are available free in full-text format. One list of such journals is available at the *Resources for Economists* webpage at <http://rfe.org>. Another list is maintained at the College of Wooster (<http://www.wooster.edu/economics/archive/journals.html>).

There are also a number of proprietary portals to full-text journals. For example, the American Economic Association hosts a web portal (<http://www.aeaweb.org>) that includes full-text access to the past three years of the journals it publishes: the *AER*, *JEL*, and *JEP*. Access to the journals’ tables of contents and abstracts is available to anyone, but full-text access requires AEA membership. The site also includes Bill Goffe’s *Resources for Economists on the Internet*, which will be discussed in detail in Chapter 9. JSTOR is an archive of full-text scholarly journals, including (as of this writing) twenty-four top journals in economics, which are listed in Table 3.5.⁹ Many colleges and universities subscribe to JSTOR. AEA members can access JSTOR via a modest annual fee (\$10). There are several other proprietary online databases that offer full-text journal articles. One of these is *Expanded Academic ASAP*, which includes the full text of many economics journals that are not included in JSTOR.

In addition to the obvious speed and convenience that online access to journal articles provides, in the near future we will be able to link directly to references cited in those articles (backward linking). It may even be possible to link to articles citing those articles (forward linking).

Government documents are also increasingly available over the Internet. In the past decade the U.S. government has made a significant effort

Table 3.5 Economics Journals Available Through JSTOR

<i>American Economic Review</i>	<i>Journal of Economic Perspectives</i>
<i>Brookings Papers on</i>	<i>Journal of Finance</i>
<i>Economic Activity</i>	<i>Journal of Human Resources</i>
<i>Canadian Journal of Economics</i>	<i>Journal of Industrial Economics</i>
<i>Econometrica</i>	<i>Journal of Money, Credit, and</i>
<i>Economica</i>	<i>Banking</i>
<i>Economic History Review</i>	<i>Journal of Political Economy</i>
<i>Economic Journal</i>	<i>Oxford Economic Papers</i>
<i>International Economic Review</i>	<i>Quarterly Journal of Economics</i>
<i>Journal of Applied Econometrics</i>	<i>Rand Journal of Economics</i>
<i>Journal of Economic History</i>	<i>Review of Economics and Statistics</i>
<i>Journal of Economic Literature</i>	<i>Review of Economic Studies</i>

to make virtually all print documents accessible online. Moreover, unlike their printed counterparts, they are obtainable for free. A good point to begin searching for government publications is at www.firstgov.gov.

The first place to look for books is still your college or university library. To date, the availability of books online is lagging behind the availability of journal articles and government documents. One very promising source for electronic books, however, is NetLibrary (www.netlibrary.com), a subsidiary of On-line Computer Library Center (OCLC). NetLibrary offers electronic texts on a subscription basis to libraries. As of this writing, NetLibrary has over fifty-four thousand titles in a variety of fields including economics. Just as you would search for a printed book, you use your college or university library's search engine to find an online text. Indeed, until you find the entry, you won't know whether the book is a hard copy or an electronic one. Once you locate the electronic copy, you can "check it out" just as you would a printed copy. Unless your library subscribes to more than one copy of the e-book, no one else from your institution has access to the book until you relinquish it. An additional benefit is that you can obtain the text twenty-four hours a day without having to physically go to the library to pick it up.

We noted earlier that libraries at smaller institutions may have only limited collections. There are several ways to get around this, such as making use of interlibrary loan services or visiting nearby research university libraries. For example, it is increasingly possible to search such library catalogs remotely to ensure that their collection has what you are looking for before visiting the library.

There are also literally tens of thousands of economics working papers available online from the National Bureau of Economic Research (www.nber.org), the Social Science Research Network (<http://papers.ssrn.com>), the Working Papers Archive at Washington University in St. Louis (<http://econwpa.wustl.edu>), and NetEc (<http://netec.wustl.edu/NetEc.html>).

SUMMARY

- In order to perform original research, one needs to determine what is currently known on a topic. Thus, at the beginning of any research project the researcher needs to survey the "literature."
- There are two sources of information on a topic: popular and scholarly. Only the latter is considered part of the literature.
- An effective search strategy includes the following steps:

- Start with secondary sources in the general topic area.
- Search for survey articles on the topic, for example, in *JEP* or *JEL*.
- Move to primary sources: for example, *EconLit* or *Social Science Citation Index*
- Use a combination of keyword searching and browsing from the references of the useful sources you've located.

NOTES

1. Appendix 10 of Wyrick (1994) provides an extensive listing of economics journals grouped by subject area.
2. Note also that working papers tend to be more readable than published articles because they usually include many more details of the argument—details that are eliminated during the editorial process.
3. This is also referred to as the *EconLit* or *Journal of Economic Literature* classification system.
4. For example, you can do online searches using these classifications as keywords.
5. The AEA website lists the tables of contents for the *Journal of Economic Literature* back to December 1994 and for the *Journal of Economic Perspectives* back to the beginning of 1998. One can easily get the contents from earlier issues of the journals by searching *EconLit*, using the journal name as the "Source" keyword.
6. Note also that you should always use multiple web search engines since they use different search criteria and their databases, while overlapping, are not identical.
7. In my courses, such a flaw would cost at least a letter grade on the final paper.
8. Note that databases like *EconLit* are available through more than one interface. For example, at my institution, we get it from OCLC FirstSearch, a database service provider that provides access to forty-four databases on a wide variety of subjects.
9. Note, however, that articles do not become available through JSTOR until five years after original publication.
10. Note that Turabian's Chapter 8 was published before the 1999 MLA decision to recommend parenthetical references.

SUGGESTIONS FOR FURTHER READING

Ackermann and Hartman (1998)—Excellent, user-friendly guide to online searches and the Internet as a research tool.

Turabian (1996)—Definitive guide to writing and citation styles for research papers and theses. See also Appendix 3A.

EXERCISES

1. Select a research question, preferably one you are actually researching. Write down the research question. Review the tables of contents for the *Journal of Economic Perspectives* and the *Journal of Economic Literature* (provided at <http://www.aeaweb.org/jep/contents> and <http://www.aeaweb.org/journal/contents>). Browse those tables of contents for one or more survey articles that are relevant to your research question. (You should identify at least one article that is relevant to your research.) Write down the complete bibliographic citation information for the survey articles you discover, using the citation style your instructor favors. (See Appendix 3A for information on citation styles.)
2. Select a research question, preferably one you are actually researching. Write down the research question. Use the basic search strategy illustrated in Table 3.4 to perform a keyword search in *EconLit*. Write down a list of “important concepts” that are relevant to your research question. Select and write down appropriate keywords. Create a search expression, and execute the search. Modify the search as necessary to find at least five useful “hits” for your research. Write down the complete bibliographic citation information for the studies you discover, using the citation style your instructor favors.
3. Select a research question, preferably one you are actually researching. Write down the research question. Use the basic search strategy illustrated in Table 3.4 to perform a keyword search using a database *other than EconLit*. Write down a list of “important concepts” relevant to your research question. Select and write down appropriate keywords. Create a search expression, and execute the search. Modify the search as necessary to find at least five useful “hits” for your research. Write down the complete bibliographic citation information for the studies you discover, using the citation style your instructor favors.
4. Select a research question, preferably one you are actually researching. Write down the research question. Use the basic search strategy illustrated in Table 3.4 to perform a keyword search using a web search engine like Google. Write down a list of “important concepts” relevant to

your research question. Select and write down appropriate keywords. Create a search expression, and execute the search. Modify the search as necessary to find at least five *scholarly* “hits” for your research. Write down the complete bibliographic citation information for the studies you discover, using the citation style your instructor favors.

Scholarly References and Citation Styles

An essential part of reviewing literature is keeping careful references for each scholarly work you use. The purpose of citation styles is to provide a standardized, concise way of doing that. References are important for two reasons: first, to give credit to previous authors for their ideas, their intellectual property, and, second, to help readers track down those ideas in full detail. We will expand on these reasons later in Chapters 5 and 6. Here we introduce the styles.

There are three general styles for citations and references: they are known as MLA, the Chicago Style, and APA. Each of these styles explains its preferred way to write scholarly notes in the text (e.g., foot- or endnotes) and scholarly references at the end of the text. All three styles now recommend using the parenthetical form for references—for example, “Turabian (1996)” —instead of citing references in the text through footnotes or endnotes. Of course, for scholarly notes that elaborate on the text, it is still appropriate to use footnotes or endnotes.

An excellent guide to the MLA and Chicago styles is Kate Turabian’s *A Manual for Writers of Term Papers, Theses and Dissertations*, Sixth Edition (1996). Ms. Turabian was the graduate secretary at the University of Chicago; she was responsible for accepting all theses and dissertations submitted at that institution and as such was the acknowledged expert on the *Chicago Manual of Style*. Thus, her book is essentially a condensed version of the *Chicago Manual*. Two chapters that are particularly useful are Chapter 11, which shows a comparison of the two citation styles, and Chapter 14, which illustrates the two styles in a large number of sample notes and references. Note that this edition of Turabian is based on the Fourteenth Edition of the *Chicago Manual of Style*, rather than the most recent Fifteenth Edition. The major changes in the new edition have to do with online resources, which were not covered in the Fourteenth Edition, so Turabian is still useful. Hopefully, a new edition of Turabian will be forthcoming. The following paragraphs introduce the three styles and refer you to various guides to each.

MLA is the acronym for the Modern Language Association and is the style favored by scholars in the humanities disciplines. This style is spelled out in Chapter 8 (on footnotes) and Chapter 9 (on bibliographies) of

Turabian (1996).¹⁰ Additional guides to the MLA style are available from the University of Wisconsin Writing Center (<http://www.wisc.edu/writing/Handbook/DocMLA.html>) and the Writer’s Workshop at the University of Illinois (<http://www.english.uiuc.edu/cws/wworkshop/MLA/bibliographymla.htm>).

APA is the acronym for the American Psychological Association and is the style used by scholars in psychology and other social sciences, including some economists. Useful APA guides are available from the On-line Writing Lab at Purdue University (http://owl.English.purdue.edu/handouts/research/r_apa.html), the University of Wisconsin Writing Center (<http://www.wisc.edu/writing/Handbook/DocAPA.html>), and the Writer’s Workshop at the University of Illinois (http://www.english.uiuc.edu/cws/wworkshop/bibliography_style_handbookapa.htm).

The Chicago style is the choice of the natural sciences and many social sciences, including economics. The Chicago Style is formally called the author-date system in the *Chicago Manual of Style*, Fifteenth Edition (2003). In this style, references are listed at the end of a work in the “Reference List” (not the bibliography). This style is explained in Chapter 10 of Turabian (1996). An additional guide to the Chicago style is available from the Long Island University library: <http://www.liu.edu/cwis/cwp/library/workshop/citchi.htm>.

Given that most literature surveys are now conducted online, you should be aware that online citation styles may not be adequately spelled out in the guides just mentioned. A good source of information about citation styles for online documents is Andrew Harnock and Eugene Kleppinger’s, *Online: A Reference Guide to Using Internet Sources*, available at <http://www.bedfordstmartins.com/online>.

Note that a journal article (or book) found online is still a journal article (or book) and should be cited accordingly. That is, online citation styles should be reserved for documents that have no print analog. For example, an article in the *American Economic Review* that you obtain online from JSTOR should be cited as a journal article rather than as an online document.

Using Writing As a Tool for Economic Research

"The important thing in science is not so much to obtain new facts as to discover new ways of thinking about them."

SIR WILLIAM BRAGG

In Chapter 1 we observed that scholars create knowledge by constructing competing arguments. They construct these arguments by using four tools:

- Mental processes—thinking about an argument;
- Oral discourse—verbalizing the argument;
- Mathematics—manipulating equations to derive meaning; and
- Writing.

Though arguments can be developed using all of those tools, in this chapter we focus on writing as the principal one. McCloskey (2000) makes this point early in her book *Economical Writing*, when she asserts that the practice of economics depends more on writing than on mathematics or statistics or any of the other technical skills normally associated with economics!

Economists use writing for two purposes, the second of which you may not have considered:

- Writing as a *Product*, a form of communication to disseminate research results, and
- Writing as a *Process* for deriving the research results.

This chapter will focus on the second purpose, while the first will be the subject of Chapter 5.

Writing to Learn

Writing is more than a product; it is also a process or a tool for creating knowledge. Students sometimes say, “I know it, but I just can’t explain it.” But if you can’t explain an economic idea, you only know it on a superficial level. Writing forces you to think concretely, to figure out exactly what you mean. By contrast, when using purely mental processes, it is easy to be vague with your thoughts, to leave pieces out. When you write, those holes in the logic become readily apparent. In this context, writing is a tool of discovery, a way of working through ideas that you don’t fully understand.¹ In other words, writing is a positive-sum game. When you write, you don’t merely put down what you already know; rather, you end up knowing more.

Composition As a Creative Process

Let’s explore this idea. The process of writing is called composition. What does that mean? Composition includes both analysis (taking something apart to understand it) and synthesis (putting pieces together to make a whole). Consider how you as a student “put together” a traditional term paper based on notes written on index cards. You have to sort through the information, classify it, and decide what is and is not relevant. You arrange and rearrange so as to discover a structure, an organization that best displays the information on the index cards until a meaning reveals itself. In other words, composition involves searching for relationships among the facts and ideas that make up the raw material of your research. When you compose, you choose and arrange those facts and ideas; you try to put them in the most meaningful order. It is the same when a director makes a film. The film is never completed in only one “take.” Rather, different takes are filmed, edited, and combined in different ways until just the right story is developed.

Students often end this composition process too early, before it fully reveals its insights. Bean (1996, 7) notes, “A key observation among teachers of critical thinking is that students . . . tend to reach closure too quickly. They do not suspend judgment, question assumptions, imagine alternative answers, play with data, enter into the spirit of opposing views, and just plain linger over questions.” This problem is understandable if one views writing only as a product. Bean (1996, 15) characterizes this mistaken view with a metaphor: “Writing is like the box and wrapping paper into which we put our already formulated ideas.”

In fact, composition is like a fine wine, which needs time to mature and develop into its full flavor. This explains why it is important to start an assignment early, write a draft, and then take time off from your writing to allow your sub-

conscious to mull over what you’ve come up with. Bean calls this “incubation.” As journeyman writers, you need to make a special effort to maintain an open mind about your topic and not allow yourself to reach a conclusion too soon.

In this light, you should view revising a paper as far more than editorial correction. It means re-viewing the information, re-thinking the way it is organized, re-constructing the argument, looking for new patterns of meaning. It means asking, “Do I have the best interpretation of the facts, or is there another that would do a better job?” This is hard work. You must take an idea, and do the best you can to build an argument for it. Then, once you’ve made a commitment to it, you need to step back and honestly scrutinize what you’ve written. This approach to revision is one reason why instructors emphasize writing multiple drafts of a paper: so you consider and explore different ways of thinking about the topic. We will discuss multiple paper drafts in more detail in Chapter 5.

One implication of this creative view of writing is that almost certainly you will have to “throw away” some of what you write, even though this is difficult to do. Even experienced writers have a sense that if they throw away parts of their writing, especially parts they have labored over, they are throwing away something of themselves. This difficulty may come in part from the misconception that term papers must include all the facts you discovered (and wrote on index cards) in the preliminary stages of the research. This is simply not so! A well-written paper is structured to create an argument. Recall the nature of a theory or a model. A model is a simplified version of reality that provides insight into the question being researched by focusing on the most important aspects of the question, while leaving out the less important aspects. In this respect, an argument is like a model; it provides a structure or story to best explain the facts, but probably doesn’t explain or address them all. This is why “throwing away” portions of your writing that upon review don’t contribute significantly to the argument is a natural part of the composition process.

The Structure of an Argument

The purpose of scholarly writing is to make an argument that is persuasive to experts in the field. Scholars use the composing process, described in the previous section, to discover and refine their arguments. When completed, scholarly writing follows a logical, hierarchical structure, in which the thesis is supported by a series of nested arguments that lead logically to the thesis as a conclusion. This is illustrated in a stylized way in Figure 4.1, which shows the pyramid structure of a well-developed student argument. You should note that actual arguments will be less mechanically structured and more individual in character.

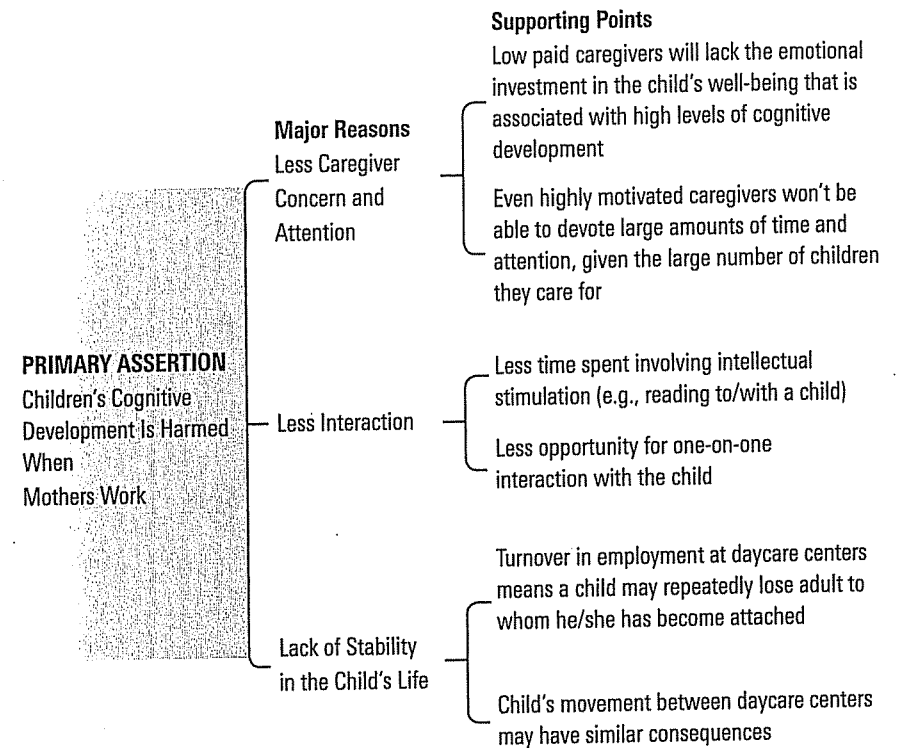
NOTES FOR NOVICE RESEARCHERS

**Composing with the Outline Function
of Your Word Processor**

The “Outline” function of your word processing software can be a very useful tool in the process of composing and revising your paper. For example, in Microsoft Word, the outline function allows you to convert your paper draft into an outline of the major points you made in the paper. The text for each section can be collapsed under each heading (or expanded to show the complete text). Once the paper is in that form, it becomes very easy to reorganize the paper without having to cut and paste. Alternatively, the outline function can allow you to create an outline of your paper that consists of the major points you plan to make, and then flesh out those points to write the complete paper. To use the outline function in Word, open your document, then select “View” from the pulldown menu on the toolbar, and then select “Outline.” Though the outline function can take a short while to master, it is well worth the effort. Indeed, I used Outline to organize and write this chapter.

The thesis or primary assertion is at the top, the major points are underneath the thesis, and supporting explanation or evidence is under each major point. Nothing is added at random. Any points (or stray facts you discovered in the preliminary research) that do not lead to the thesis either directly or indirectly are omitted from the argument.

Let’s examine this structure in more detail by looking at the nature of an argument. In Chapter 1, we defined an argument as an assertion supported by reasoning or evidence. An assertion is a claim or point of view. For example, McCloskey’s statement mentioned at the beginning of this chapter is an assertion. In casual conversation, people often argue on the basis of assertions alone (what Missimer [1995] calls “loose arguments”). Similarly, in professional or scholarly writing authors might refer to McCloskey’s “argument” that writing is more important to the work of economists than technical skills. But argument in the strictest sense refers to both the assertion and the supporting reasoning.²

Figure 4.1 Pyramid Structure of an Argument**What Does It Mean to Say That a Conclusion Follows from the Evidence?**

For an argument to be persuasive the reasons supporting it must be true, and the conclusion must follow from those reasons. What does it mean to say that a conclusion *follows* from the evidence? An **inference** is a conclusion reached after reasoning logically about facts and relationships. If we define an argument as a claim supported by evidence, then an inference is like an argument turned backward. In other words, given certain facts or assumptions, given certain relationships between the facts, if we can reason our way to the inference that is, if the evidence leads us logically to the inference as a conclusion, then we say that the conclusion “follows.” If the conclusion follows, then we have what Missimer calls a “warranted inference.”

Consider, for example, the argument that consumer spending increased last year because of an increase in the unemployment rate. To assess this argument, one would ask two questions. First, is the reason true?; did unemployment increase last year? Suppose the answer is yes. The second question then is: Does the conclusion follow from the premise? In other words, is it reasonable or logical to conclude that an increase in unemployment would cause individuals to spend more? Probably not! An increase in unemployment would reduce incomes, at least for the unemployed. Lower incomes should translate into less spending rather than more. In this case, we say that the conclusion doesn't follow from the premise, and the argument fails to convince. By contrast, if the argument stated that consumer spending increased because of a decrease in unemployment, then the conclusion would follow, and assuming unemployment had fallen, the inference would be warranted.

A **logical fallacy** is an argument that is flawed because the conclusion does not actually follow from the reasons stated, even though the argument is phrased in a way that makes you think it does. Logical fallacies often find their way into public arguments. Appendix 4A lists the most common logical fallacies, which you should learn to identify and avoid.

Examining an Argument

Let's examine the argument made by Robert Samuelson (2002), which is reproduced in the following box.

A War We Can Afford

By Robert J. Samuelson

Wednesday, September 18, 2002; Page A29

A possible war with Iraq raises many unknowns, but "can we afford it?" is not one of them. People inevitably ask that question, forgetting that the United States has become so wealthy it can wage war almost with pocket change. A war with Iraq would probably cost less than 1 percent of national income (gross domestic product). Americans have grown accustomed to fighting with little economic upset and sacrifice.

Samuelson's Major Claim: A possible war with Iraq raises many unknowns, but "can we afford it?" is not one of them.

The last time the United States truly mobilized for conflict was World War II. Roughly 16 million Americans served in the military; that was two-thirds of all men from 18 to 34, reports historian James Patterson of Brown University. The costs were stupendous. In 1944, federal spending totaled 44 percent of GDP, with military spending at 38 percent of GDP. At home, Americans needed ration coupons to buy meat, gasoline, and other staples.

Ever since, two things have transformed the economics of war: The U.S. economy has gotten bigger, and wars have gotten smaller. Measured by what it produces—and adjusted for inflation—the economy is more than five times as large as it was in 1945. Meanwhile, America's wars have become more localized, draining less of the nation's wealth.

In the Korean War, the defense budget reached 14 percent of GDP in 1953, but much of that spending went for a massive buildup of forces in Europe. "For every tank that went to Korea, two went to Europe," says Historian Alan Gropman of the Industrial College of the Armed Forces. "The B-52s we built had nothing to do with the Korean War." In August 1949, the Soviet Union unexpectedly exploded its first atomic bomb, prompting President Harry Truman to order the National Security Council to undertake a major review of U.S. strategy. The resulting document (called NSC 68) envisioned huge deployments of U.S. troops to Europe to counter a conventional Soviet attack, which—once the Soviets had their own nuclear weapons—seemed more credible.

Although the Cold War's costs remained large, defense spending during the Vietnam War went only as high as 9.4 percent of GDP, in 1968. Even so, Lyndon Johnson's early attempt to finance the war without any tax increase—to have both guns and butter—helped raise inflation. After Vietnam, defense spending (again, as a share of GDP) drifted down and dropped sharply once the Cold War ended.

It now runs about \$350 billion annually. That's a lot of money, but, in an economy producing more than \$10 trillion annually, it isn't much of a burden. It's slightly more than 3 percent of GDP. How much a war with Iraq would cost is guesswork. It would vastly exceed the total in Afghanistan,

Major Piece of the Evidence: We can measure the financial burden of a war on a nation by the cost of the war as a share of gross domestic product (GDP). During World War II, the cost of the war reached nearly 40 percent of GDP.

Crux of the Argument: Since World War II, wars have become more localized and less costly, while the economy has grown substantially.

Supporting Point: Discussion of the Korean War (more localized and less costly than World War II)

Supporting Point: Discussion of the Vietnam War (more localized and less costly)

which the Congressional Budget Office estimated at \$10 billion for fiscal 2002. The Persian Gulf War cost \$61 billion. Even if a new war cost \$100 billion, it would be only about 1 percent of GDP.

Clearly affordable. Whether we should afford it is another question. So is the effect of a war on the economy, which could go either way. Extra spending might help. A swift victory might bolster confidence. Or a war might jeopardize oil supplies from Saudi Arabia and elsewhere, through terrorism or political upheaval. Protracted fighting might hurt confidence, consumer spending and the stock market.

"The economy is growing, but barely," says Mark Zandi of Economy.com. "It wouldn't take too much for the recovery to roll back into recession—or close to it."

Oil prices are the biggest vulnerability. In 1990, after Saddam Hussein invaded Kuwait, prices doubled, from about \$18 a barrel in May to \$36 in October. That helped tip a weak U.S. economy into recession—though prices quickly receded in 1991. Barring a calamity, many economists minimize the odds of a repetition.

If Iraq's oil exports "are lost, it wouldn't be a big factor," says John Lichtblau of the Petroleum Industry Research Foundation. Iraq is now exporting about 700,000 barrels a day out of total world demand of almost 77 million barrels daily. The Saudis could offset any shortfall, he says. Moreover, the U.S. economy has become less energy-intensive and is less sensitive to higher prices. The economy is growing at an underlying rate of 2 percent to 2.5 percent, says economist Nariman Behraves of the forecasting firm DRI-WEFA. War fears have already pushed oil prices to about \$30 a barrel, but even if they rose to \$40 for a year, GDP growth might drop only half a percentage point, he says.

Because wars surprise, who knows? But the important questions are harder. Is this war justifiable? Should the United States go it alone? What will happen if we don't fight? What will happen if we do? By contrast, economic issues are a sideshow—and should stay so. If this war is necessary, we can afford it.

Copyright © 2002, *Newsweek*. Reprinted with permission.

To summarize Samuelson's argument:

Assuming that we can measure the financial burden of a war on a nation by its cost as a share of GDP, then since World War II, U.S. GDP has increased dramatically. At the same time, wars have become more localized and thus less costly; therefore, we can conclude that the United States is better able to afford a war with Iraq than we were able to afford World War II.

Samuelson's conclusion does follow from the evidence.

Three Types of Reasoning: Deductive, Inductive, and Warrant-Based

Every argument consists of assertions and evidence. Or, looking at it backward, every argument consists of premises (or assumptions) and conclusions (or inferences). Let's examine more carefully the ways in which scholars reason. There are three different approaches to reasoning: deduction, induction, and warrant-based. An argument can include all three types, but to keep things simple let's look at them one at a time.

Deductive reasoning starts from one or more general principles and derives specific predictions from them. These predictions are what Sherlock Holmes called "deductions." The principles need not be all that weighty; you can think of them as premises or assumptions. They are the "givens" in the exercise. A **valid deduction** is one in which the conclusion *must follow* from the premises. In other words, if the premises from which the reasoning begins are true, then the conclusion must be true also. In a sense, a deduction is a logical proof. Consider a simple example: Start with the premise that all monetarists believe in some version of the quantity theory of money. If Milton Friedman is a monetarist, then he must believe in the quantity theory of money. The deduction is that Friedman believes in the quantity theory of money. Samuelson's argument in the article we examined previously was another example of deductive reasoning.

Note that, strictly speaking, a deduction is considered valid as long as the conclusion follows from the premises, *even if the premises aren't true*. To be persuasive, what we really want is a **sound argument**, that is, a valid deduction in which all the premises are true.

When scholars "theorize," they are typically using deductive reasoning. For example, consider the origins of the law of demand. Assume that individuals derive satisfaction or utility from consuming goods and services,

but that their consumption is limited by each individual's budget. So the sum of each individual's expenditures on all goods purchased must not exceed the individual's income. Assume further that the utility individuals derive from consumption follows certain general principles: that increased consumption of any good or service yields more utility, but at a decreasing rate. Thus, consumers are subject to the law of diminishing marginal utility. Finally, assume that individuals wish to spend their income on the combination of goods and services that maximizes their (total) utility. It can be shown by manipulating these assumptions that if the price of a good or service rises, the amount of that good or service that individuals will purchase should fall. This is, of course, the law of demand. The process of theorizing will be addressed more fully in Chapter 7.

An **inductive argument** is one that reasons in the opposite direction from deduction. Given some specific cases or situations, what can be inferred about the underlying general rule? The reasoning *process* follows the same steps as in deduction: one reasons from one or more premises to a conclusion. The difference is in whether the premises or the conclusion is the general principle or the specific case. An inductive argument is not a proof in the sense that a deduction is. Rather, induction is a probabilistic inference. If the truth of the premises increases the likelihood that the conclusion is true, then we have an inductive inference. It is still possible for the rule to be false, however, despite the cases that lead one to believe that it is true.

When scholars use statistical evidence as proof of a hypothesis, they are using inductive logic. Suppose a researcher believes that the price of the product affects the amount the firm chooses to produce. The researcher collects a sample of data, runs appropriate statistical tests, and discovers that based on the data sample, an increase in the price of the product corresponds to an increase in the quantity supplied by the firm. We call this general rule the law of supply, which in this example was validated by a specific data set. In short, the specific data resulted in the general rule. That is induction. The use of empirical methods to test hypotheses will be explained in detail in Chapters 10 and 11.

Missimer (1995) suggests that we can learn more by focusing on the logic of arguments than by labeling the reasoning as inductive or deductive. This suggestion is especially true, since complex arguments can include both types of reasoning, as well as the third type: warrant-based.

Warrant-based reasoning is often used in scholarly writing. **Warrants** are unstated or underlying assumptions on which an argument stands. Often warrants are higher-order assumptions or axioms that are not testable. For

example, in the previous example that deduced the law of demand, we assumed that consumers maximize utility. This is a generally accepted assumption in economics, but if you think about it, how would one test that? How can we be sure that consumers try to maximize their utility rather than settling for, say, 90 percent of their potential utility? (How can we be sure that students try to maximize their GPA, rather than pursuing some other objective in school?) In fact, we can't be sure.

The purpose of a warrant is to establish the relevance of the evidence in supporting some claim. Consider the following claim:

An increase in the excise tax on beer would decrease beer consumption among underage drinkers.

What is the connection between a tax and beer drinking? What links the reason to the assertion? What makes the reason acceptable as evidence? The warrant here is the law of demand, something every economist is familiar with: increasing the tax would effectively raise the price of beer, thereby decreasing the quantity of beer consumed.

When scholars write to an audience of their peers they tend not to explicitly identify their warrants. This is because a research community is defined in part by the warrants it shares. As a result it is unnecessary for members to state them outright. Additionally, this practice allows scholars to write their arguments in a compact form. As a novice researcher, until you become familiar with the warrants in a research community, you should explicitly state all the warrants in your argument.

Complex arguments, such as those found in scholarly articles and books, can include all three types of reasoning. For example, the theoretical analysis may combine elements of deductive and warrant-based reasoning, while the empirical analysis will typically be inductive reasoning.

What Makes for a Persuasive Argument?

So, where does that leave us? We noted earlier that for an argument to be persuasive, the reasons supporting it must be true, and the conclusion must follow from the reasons. The evidence of the strongest arguments, however, has a number of additional complementary characteristics.³ In particular, the evidence should be accurate, authoritative, precise, clearly explained, complete, and representative. Let's turn to each of these characteristics.

If the supporting reasons must be true, then *factual evidence needs to be accurate*. If your argument depends on unemployment having risen, then you need to be sure that it did. If you don't check to confirm that it did,

some reader will and may find that it didn't. Similarly, when you analyze data, a key part of empirical research, you need to confirm that your data are correct. Something as trivial as a data-entry error can completely change the results of a statistical analysis.

Evidence should also be authoritative. If you are using data, is it from a reputable source? A student of mine once investigated the demand for Beanie Babies. She concluded that the demand for Beanie Babies as a collectible should depend on their expected future price as well as the current price, the buyer's current income, and other items one might expect to see in a demand function. She obtained data for expected future prices, which her statistical analysis revealed to be the most important determinant of demand for Beanie Babies. Unfortunately, she obtained the data from the manufacturer, which was not an objective source, since it stood to benefit significantly if its products were seen to be a good investment.

Similarly, if you support an assertion with a quotation, is the quotation from an authoritative source? If you cite an argument, is it from a respected publication? In Chapter 3, we discussed the way in which scholarly publications are vetted, making them authoritative. The Internet provides wonderful examples of the pitfalls of using evidence that lacks authority. Since virtually anyone can publish virtually anything on the Internet, one needs to be extremely cautious about basing conclusions on evidence discovered on the Internet. Booth et al. (1995, 102) declare that if you are not an expert in the field, "do not trust any source as authoritative until you know the research in the area."

Additionally, *evidence needs to be precise.* Arguing that welfare reform is a bad policy because *it harmed many people* is weaker evidence than arguing that 3.7 million people lost benefits. Imprecision or outright vagueness will weaken the reader's perception of evidence, just as mistakes or carelessness do. Imagine trying to hit a baseball with a Nerf bat versus a wooden one. The Nerf bat will diffuse the energy of your swing, so it will be hard to hit the ball with any force. In the same way, a vague reason lacks power. (Note, however, that there is such a thing as too much precision, for example, reporting regression coefficients to eleven decimal points.)

Evidence is rarely self-evident. One nearly always needs to explain it. This is especially true of statistical results. Undergraduate researchers sometimes believe that statistical results must merely be presented to be convincing, but as we will explain in Chapter 12, statistics must be interpreted for the reader. If the evidence is not *clearly explained*, readers may not see how the conclusion follows from the reasoning. If a reason is not clear to the reader, it *will not* help persuade him or her that the argument

is true. In fact, it may do the opposite. Booth et al. (1995) emphasize that it is often helpful to ask the question: Will the reader see the evidence as you see it? Is there additional information that you have that establishes the relevance of the evidence? If so, you may need to provide it.

To construct the strongest possible case, *the evidence needs to be complete*, that is, it should have depth and breadth. You can have a sound, logical argument *but still not persuade* if your argument fails to consider all the relevant evidence (pro and con) or doesn't address it in enough detail.

Deep reasoning requires you to think deeply about causes and effects. Why might unemployment have risen? A student might respond, "Perhaps because businesses are producing less output, so they require fewer workers." There is nothing wrong with that reasoning, except that it doesn't go far enough. If that was the extent of an answer to an exam question, it might be labeled a "shallow argument," one that lacks depth. A better answer would go on to ask, why are businesses producing less output? The answer could be "Because sales are off." Why then have sales fallen? Perhaps it is due to the decline in the stock market, which makes consumers feel less wealthy and businesses feel pessimistic about the future. When you reason deeply, you follow the train of cause and effect as far as you can.

In a similar way, broad reasoning requires you to think widely about causes and effects. Instead of being satisfied with just one or two reasons for an assertion, you should try to identify all possible reasons. Booth et al. (1995) suggest that you try to anticipate all the questions a thoughtful reader might have about your argument, and then address those questions in your argument.

The evidence presented in your argument should also be *representative* of thought on an assertion. For example, students sometimes think they should only report the evidence that supports their assertion. Nothing could be further from the truth! Aside from being intellectually dishonest, this opens your argument to the correct criticism that you haven't considered the opposing viewpoint.⁴ Failure to do so will lead knowledgeable readers to label you ignorant at best, and fraudulent at worst. You are not trying to win a debate, but rather determine the best answer to the question. Moreover, as Booth et al. (1995) point out, one can actually improve the strength of one's argument by admitting its limitations and conceding contrary evidence. In a statistical context, the same issue arises: one's sample of data should be representative of the underlying population if inferences based on the sample are to be valid. This important point will be discussed in detail later in Chapter 10.

Example: Evaluating an Argument with Unrepresentative Evidence

Suppose the average GPA of students who graduate from college is a B. Does that imply that grade inflation is occurring? After all, if there were no grade inflation, shouldn't the average grade be a C? There are at least two reasons why this may not be a warranted inference. The first involves the GPA requirement for graduation. Most schools require seniors to have at least a C average to graduate. Thus, even if the average GPA of the senior class is a C, the GPA of *graduates* must be higher, since those seniors with a lower GPA are by definition excluded from graduation. In other words, the sample of data used in this argument (i.e., seniors who graduate) is not representative of the population of students.

The second reason the conclusion does not follow in this argument involves the courses students take over their undergraduate careers. Most students probably end up taking courses in and perhaps majoring in subjects in which they do well, or at least better than average. This implies that almost certainly the GPA of the "average" graduate will be higher than a C. Again, the sample of data used is probably not representative. A better test might be to examine the grades students that earn in their general education courses, which every student must take, and especially those taken in their first or second year of college.

Example: Evaluating an Argument Using Broad and Deep Reasoning

How can one explain the increased purchases of sport utility vehicles by Americans during the late 1990s? One answer might be the increase in personal incomes over that period. But is that the only possible reason? After all, real personal disposable income rose only 28 percent while sales of SUVs roughly tripled. The law of demand suggests an inverse relation between price and quantity demanded. If the price of SUVs fell, that could also explain the trend. But in fact SUV prices rose, so that reason isn't valid. What other reasons might be suggested by demand theory? Could it be a decrease in the price of gasoline, a complementary good? What happened to the price of minivans, a substitute good? When you reason broadly, you consider as many reasons as you can. Inexperienced critical thinkers need to make a special effort to cast their net widely before drawing conclusions to ensure they don't miss one or more relevant reasons.

How broadly and how deeply must one think? One rule of thumb is to consider the weightiness of the primary assertion one is arguing. A controversial thesis with large implications will require broader and deeper explanations than a more commonplace assertion. Consider the recent

argument that the information age has fundamentally changed the workings of the U.S. economy with the result that U.S. productivity growth has significantly increased. This assertion, if true, will have major implications for the standard of living in the United States, as well as for macroeconomic policy. Though a great deal of evidence supports this thesis, the weight of its implications means that for most analysts, it remains an open question.

Example: Constructing a Simple Argument

Consider the following assertion: "At the next meeting of the Federal Open Market Committee, the Federal Reserve will almost certainly decide to lower interest rates again."

What might be some reasons in support of this assertion? What might lead the Fed to pursue a more expansionary monetary policy?

- The unemployment rate is the highest it has been in four years.
- Consumer confidence is falling.
- Business investment spending is down.
- Inflation is negligible.

What might be some underlying reasons that support these larger reasons? For example, what might explain the high unemployment rate or the falling consumer confidence; what is the implication of low inflation?

- U.S. production has slowed substantially in the wake of the stock market decline over the past year.
- Households are worried about the future.
- Given the decline in U.S. production, the capacity utilization rate is down.
- The lack of inflation allows the Fed to pursue a more expansionary policy than it might otherwise be comfortable with.

Notice how the principal assertion and the major and minor reasons form the pyramid structure of the argument, as illustrated in Figure 4.1. Assuming that the assertion follows from the reasons and that the reasons are valid, then the more complete the reasoning, the more solid will be the base of the pyramid and the more solid the argument.

Note also that when you begin to construct an argument, you may not know which way you will end up arguing. Rather, you should start with the research question. Next, you consider all possible answers pro and con. Next, you consider all the reasons for the answers. Finally, you evaluate the evidence and make a judgment. This judgment becomes the assertion, which is supported by your reasoning—that is, your explanation of how you evaluated the evidence and why you came to the decision you did.

An Important Caveat

A conclusion can follow from the evidence, the evidence can be correct, and the argument may still be incorrect. In other words, internal consistency in an argument is necessary but not sufficient if another conclusion explains the evidence better and more correctly. In the real world, scholars rarely have all of the relevant evidence. Rather, they construct their arguments based on a sample of the evidence that they think is the most relevant. It is always possible that they will miss some critical piece of evidence (an assumption, a piece of data, a relationship between parts of the evidence). As a result, their argument will be incorrect.

The way to minimize this possibility is to insure that your argument has the features outlined in the last section, including broad and deep thinking. But you can never be certain. Therefore, a modicum of humility is always in order when presenting a scholarly argument, even a good one.

SUMMARY

- Writing is a tool for creating arguments.
- When you develop an argument you need to think deeply and broadly about your reasoning.
- Each conclusion should follow from its premises.
- Can any assertions be made more convincing with more elaboration?
- Are there other reasons to support the conclusion?
- Is an alternative explanation more plausible?

NOTES

1. Knoblach and Brannon (1983). Zinnser (1988), one of the founders of the Writing to Learn concept, observes, "We write to find out what we know." McCloskey (2000) similarly states, "Writing is thinking."
2. Booth et al. (1995) define an argument as including an **assertion** (or claim), **evidence** to support the claim, a **warrant** (or general principle that explains why you think your evidence is relevant), and **qualifications** (or limits on the assertion). The last three elements are what we referred to as "supporting reasoning" in our definition of "argument."
3. This section draws heavily from Booth et al. (1995).
4. This is called the fallacy of special pleading. See Appendix 4A.
5. Robert Weissman, "Why we protest," *Washington Post*, September 10, 2001, p. A21.

SUGGESTIONS FOR FURTHER READING

- Booth et al. (1995)*—Classic text on college-level research. Chapters 7–10 explain in detail how to construct scholarly arguments, but at a level that is very accessible to undergraduates.
- Graff (2003)*—Thoughtful critique of academia as needlessly obscuring what education is: being able to understand the arguments of others as well as to construct one's own. See, especially, Chapters 4 and 5.
- Epstein and Kernberger (2005)*—High-level but very readable guide to the logic of economic analysis. Examples are taken from classic economic literature and current texts.
- McCloskey (2000)*—The original guide to writing in economics; first edition published in 1987. Very good, but also very opinionated. Not every point McCloskey makes about composition would be considered law by writing instructors.
- Missimer (1995)*—Excellent introduction to critical thinking and how to build written arguments.
- Trelogan (2001)*—Useful introduction to logic for undergraduates.

EXERCISES

1. Consider the following excerpt from the *Washington Post*:⁵

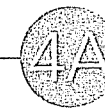
The IMF and World Bank must cancel the debts owed them by impoverished countries. Poor countries, including those that have passed through the institutions' debt relief program, routinely spend more money servicing foreign debt than they do on health care or education. U.N. Secretary General Kofi Annan has called on rich countries to provide \$10 billion a year to assist poor countries ravaged by AIDS, tuberculosis, and malaria. By what logic should the people of these same countries be forced to transfer money to rich countries? The IMF and World Bank have sufficient funds and assets in their coffers to undertake debt cancellation without additional money from U.S. taxpayers.

What is the principal assertion here? What are the reasons given? What are the underlying assumptions? Are the reasons valid? Does the conclusion follow from the reasons? Do you find the argument persuasive? Why or why not?

2. Construct an argument to support the following assertion: The next quarter's GDP growth figures will show that the United States has begun an economic recovery. Identify several major reasons in support of this

assertion, and at least one minor reason to support one of your major reasons. Write up your answer in the pyramid structure of a good argument (i.e., Main Assertion at the top of the page, Major Reasons below, Minor Reasons below each Major Reason).

3. Divide the class into two groups: Each group should construct an argument about the need for tax cuts, one group in support, and one group opposed. Each group should try to anticipate the evidence of the other side and include/rebut it in their arguments. Bring the two arguments together to make a meta-argument, and have the class, acting as individuals, decide the conclusion.



Logical Fallacies

A logical fallacy is an argument that is flawed because the conclusion does not actually follow from the reasons stated, even though the argument is phrased in a way that makes you think the conclusion follows and even though the reasoning has appeal. Logical fallacies often find their way into public arguments since they can be persuasive to many people. Here is a list of some commonly used logical fallacies:

- **Straw Man**—Mischaracterizing a position by omitting its strongest reasoning, often by caricaturing it in a way that no one would agree with. Thus, one is not addressing the actual argument posed. For example, “Alan Greenspan is opposed to budget deficits because he is concerned that the U.S. government might go bankrupt.”

(Greenspan has consistently opposed budget deficits on the grounds that deficits increase real interest rates, which diminish capital investment and ultimately reduce per capita GDP. The extent to which this occurs is an empirical question about which reasonable economists disagree.)

- **Special Pleading**—Selectively using the available evidence; only using evidence that supports your position; ignoring any evidence in opposition. For example, “We should run a budget deficit since it will stimulate the economy.”

(Yes, it will in the short run, but it may also adversely affect the economy in the long run.)

- **Begging the Question**—Making an assertion in which the reason given doesn’t really support the conclusion; hence, a non sequitur. Often this appears as a type of circular argument, where the “reason” is only a restatement of the assertion. For example, “Consumers buy less at a higher price because of the law of demand!”

(The law of demand can be paraphrased as an inverse relation between price and the amount consumers buy. Thus, the statement doesn’t explain why consumers buy less.)

- **Affirming the Consequent**—Drawing conclusions based on unexamined premises. For example, “Government fine-tuning of aggregate demand would result in a more stable macro economy.”

(Theoretically, yes, but this assumes that government is capable of fine-tuning—an assumption that many macro economists question.)

- **Ad Hominem**—Refuting an argument by attacking the person, rather than his or her argument; rejecting an assertion because of who is making it rather than because of the evidence. For example, “His argument about the budget deficit must be wrong because he’s a liberal”; “Her argument about tax cuts must be wrong because she’s a conservative.”

(An argument should be evaluated on the basis of its logic. Note that this is somewhat the opposite of an appeal to authority!)

- **Appeal to Authority**—Accepting an argument because an expert endorses it. (Think about this fallacy the next time your instructor makes an assertion!) The “expert” may have no particular expertise on this issue. Alternatively, this may be a question on which different experts disagree. The argument may be correct, but it is a fallacy to accept it without examining the logic and evidence. For example, “Budget deficits should not be a concern since President Bush is not worried about them.”

(Bush may not be worried about them, but he is a political leader rather than an economist! Some economists have a different view of deficits.)

- **Appeal to the People/Appeal to the Many**—Accepting (or rejecting) a position because many others do, again without examining the argument. For example, textbook sellers often tell faculty, “You should use this book in your class because many other schools do.”

(My course is not the same as the courses at all those other schools; my students are not like the students at all those other schools.)

- **Post Hoc, Ergo Propter Hoc** (literally, “After this, therefore because of this”)—What comes before was the cause. Though it is intuitive to think that what comes before causes what comes after (e.g., a college degree leads to a good career), it is not necessarily true. In fact, the cause may come *after* the effect. For example, does the buildup of toy inventories in the first three quarters of every year “cause” Christmas? This is a variation of the important dictum that correlation is not necessarily causation. For example, in a cleverly titled study invoking this fallacy, James Tobin asked “Does the Fed cause Christmas?”

(Even though the money supply tends to grow in the fourth quarter of every year, this is in anticipation of Christmas rather than a cause of it.)

- **Fallacy of Composition/Fallacy of Division**—What’s true at the micro level must be true at the macro level, and vice versa. For example, the classical economists argued that the solution to the widespread unemployment of the Great Depression was to decrease real wages.

(Keynes pointed out that if all workers face lower incomes, they would reduce their consumption, further lowering the demand for goods and thus for labor.)

- **Appeal to Pity**—Using sympathy for one issue as justification for another issue. For example, “My research paper deserves a good grade because I’ve been taking eighteen credit hours this semester.”

(It is impressive that you’ve been taking so many courses this term, but the grade you get on your research paper should be based on the quality of the paper.)

- **False Analogy**—Drawing parallels between two cases where there are enough substantive differences to question the comparison. For example, “The way to balance the government budget in Argentina is to raise the income tax rate.”

(Though such a policy could work in a developed country where people pay their taxes, in developing countries like Argentina there is widespread tax avoidance, so merely raising the rate may not work. Indeed, it might have the opposite effect because of a Laffer Curve effect.)

A comprehensive discussion of logical fallacies can be found in Stephen Downes, *Stephen’s Guide to the Logical Fallacies*, <http://www.datanation.com/fallacies>.

Writing As a Product of Economic Analysis

*"Planning to write is not writing.
Outlining . . . researching . . . talking to
people about what you're doing, none of
that is writing. Writing is writing."*

E. L. DOCTOROW

In Chapter 4, we explained how the writing process could be used as a tool for developing one's argument. In this chapter we discuss the use of writing as a product of the intellectual process.

Writing as a product is different in several ways from writing as a process. Writing as a process is a tool the author uses to determine what he or she thinks about a question. As such, issues like format, style, and grammar are largely irrelevant, since you are writing to yourself. Writing as a product is the report of that inquiry. Thus, the audience is someone other than the author. We observed earlier that scholarly writing embodies an argument that attempts to persuade experts in the field. Thus, the writing needs to be more explicit and formal. This means that everything needs to be spelled out clearly and in sufficient detail to get your point across. Additionally, you must follow proper standards of punctuation and grammar.

What Is Economic Writing?

Students often tell me that they don't know how to write an economics paper, or that they don't know what makes a paper economic in character. Palmmini (1996) identifies at least four different types of writing done by economists: Research Studies (including senior theses, dissertations, journal articles, monographs, and scholarly books); Social Criticism (including

op-ed pieces and books for educated lay audiences); Policy Analyses (including technical reviews of legislative or regulatory proposals and other government reports); and Business Analyses (forecasts, market analyses, cost analyses, and similar analyses).

Though many features of writing are common across all disciplines, certain aspects are discipline-specific. Petr (1998, 229–230) points out that “Good academic writing . . . rel[ies] on a unique blend of vocabulary, concept, method, precedent and history, typifying that discipline and reflecting its thought-patterns (its metaphors, if you will).” Good disciplinary writing requires more than the ability to cut and paste phrases together from the vocabulary of the field. What is required is a higher order of cognition: understanding the appropriate disciplinary context for those phrases. Bean (1996) observes that each discipline asks certain types of questions and has its own way of analyzing them. Each discipline uses its own types of arguments, proofs, and empirical evidence.¹

What fundamentally distinguishes economic from other disciplinary writing and what all types of economic writing share is the use of economic analysis. We introduced this concept in Chapter 2 when we discussed what makes for a research question in economics. In other words, all economic writing applies economic theory to derive insights about and explain answers to a question or problem. The types of theoretical arguments and empirical evidence economists use will be explored in Chapters 7, 10, and 11.

Writing Steps

Writing experts have identified four steps necessary to complete a finished paper:

1. **Pre-writing or Exploration**
2. **Writing the First Draft**
3. **Revising**
4. **Editing**

The first step was the subject of the last chapter. The last three will be the subject of this one. Though we describe this process as four steps, writing, like research more generally, isn’t usually a linear process. Instead, it’s iterative; many writers loop through the steps more than once. Thus, you can expect to read about many of the same issues in the later section on “Revising” that you do in the following section, “Writing the First Draft.”

Writing the First Draft

One of the hardest steps in completing a paper is writing the first draft. Even professional writers find this to be the case. How do you start this

NOTES FOR NOVICE RESEARCHERS

Allocating Your Time Between the Writing Steps

Inexperienced writers spend most of their time and effort on drafting, relatively less on pre-writing, and relatively little on revising and editing. Experienced writers reverse those proportions. They spend the most time on pre-writing and revising and the least time on drafting. You would be well served to follow the practice of experienced writers.

step? You can prepare for the first draft by writing throughout the research process: taking notes as you research your topic, writing critiques, drafting your own ideas (Booth et al., 1995, 149).

The first draft really begins in the pre-writing or exploration stage. It is at this stage that the writer identifies and develops the argument to be made in the paper. This is done, as we described in Chapter 4, by reviewing and attempting to organize your research materials.

Before you start writing, there is one key question you need to address: who is the audience for your paper? We noted at the beginning of that chapter that one significant difference between writing as a process and writing as a product is the audience. We also noted that this is not a trivial question since a research paper really has two audiences: the scholarly community for your topic and the instructor and classmates who are likely to read your paper. It is a good idea to ask your instructor who your audience should be. I encourage my students to write to economics majors who are at the same level as they are. In any case, this question is not one you should ignore. As McCloskey (2000) observes, if you fail to identify your audience, your paper will almost certainly miss the target.

Features of Good Economic Writing

Once you have your argument sketched out, you are ready to write. What are you aiming for? What does a well-written paper look like?

Good economic writing has several features. These are illustrated in order of importance in Figure 5.1. First, good writing should be *focused*, not fuzzy. Early in your paper, you should make clear to the reader exactly what the paper is about. You can do this by inserting a thesis sentence that describes the principal assertion of the paper. Experts differ on what construction this sentence should take. Wyrick (1994) suggests, “The purpose

Figure 5.1 Features of Good Economic Writing

1. Good writing should be *focused*, not fuzzy;
2. Good writing should be *organized*;
3. Good writing should be *solidly developed*;
4. Good writing should be *clear, concise, and precise*;
5. Good writing should be *free of grammatical errors*.

of this paper is . . .” McCloskey (2000) considers this poor style. Perhaps it is, but my advice to writers is to first communicate clearly; then worry about style. If you need to use Wyrick’s construction to get your point across, that’s fine. As you become a more experienced writer, you’ll be able to get beyond this.

Figure 5.2 contains passages from two undergraduate papers. Read each one and see if you can tell which one is more focused.

What is the primary difference between the two passages? Can you identify the thesis sentence for the passage on the left? Can you do it for the passage on the right?

Most readers conclude that the passage on the right is more focused, primarily because of the clear thesis sentence at the end.

The second feature of good economic writing is that it should be **organized**. This is far more than a question of format (e.g., introduction, body, conclusion). Rather, scholarly papers should follow a logical, hierarchical structure in which the thesis is supported by a series of nested arguments that lead logically to the thesis as a conclusion. This was the subject of Chapter 4. Recall the pyramid structure of a well-developed argument illustrated in Figure 4.1? The structure of an argument may be a pyramid, but a written document is by its nature a linear structure. We can see this in Figure 5.3, which converts the argument from Figure 4.1 into the linear organization of a paper.

To determine the organization for your paper, you should ask several questions: What are the major points that you want to make that lead logically to your conclusion (i.e., that prove your thesis)? These points will be the principal pieces of evidence to support your argument. What is the best order for those points in the paper? Are there any weak or missing links in the logic? If so, fix them. Are there any superfluous points (i.e., points that don’t lead to the conclusion)? If so, omit them!

These guidelines for an argument are true regardless of its length. In a short essay of one hundred words, the thesis is likely to be either at the be-

Figure 5.2 Examples of Focus

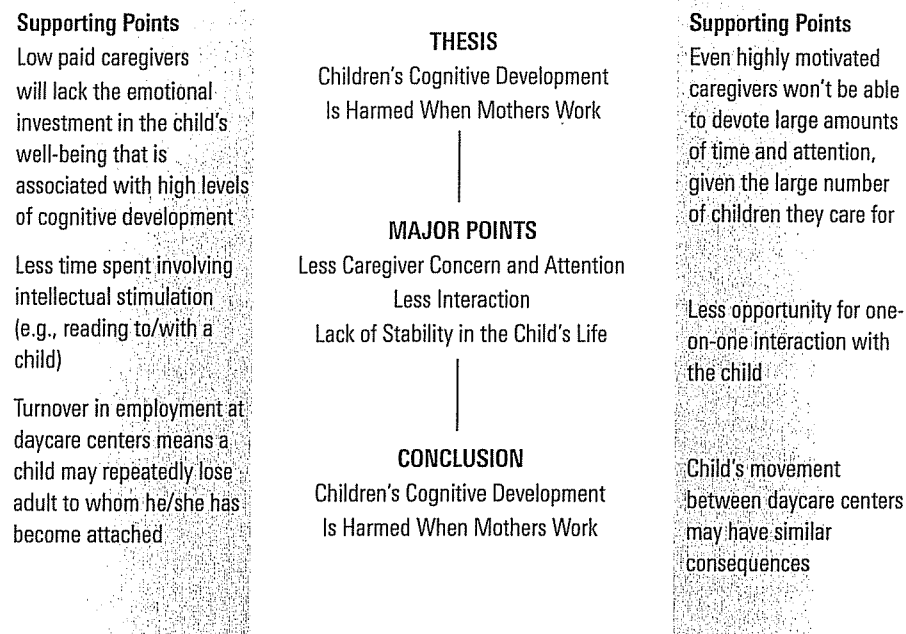
In the postcolonial period, several new African nation-states experimented with economic development ideologies. Tanzania began to ostensibly establish regional hegemony in the postcolonial period, under the auspices of the socialist world, immediately after independence in 1963. The Arusha Declaration of 1967 formalized the state’s ideological orientation toward “Ujamaa,” the collective (state-owned and centrally controlled) ownership of state-granted monopolies (Bagachwa, 1990). This declaration, however politically and rhetorically laudable, did little to increase the productive capacity of the country. With a quickly decaying infrastructure, nationalistic development policies, high levels of regional political and investment risk, numerous investment complications and capital flight, Tanzania’s socialist legacy has left a microcosmic wake of typological economic development problems for the world’s less developed countries (LDCs) (Bagachwa, 1990).

Growing concerns over the increasing juvenile crime rate have sparked extensive research into the socioeconomic and psychological factors that lead young adults to engage in illegal behavior. Over the past forty years, the number of juvenile court cases handled has more than doubled. The rise in destructive acts among youths is evident in school dropout rates, teenage pregnancies, drug abuse, and quite apparently, in the substantial increase in teenage crimes committed. Over the past several years, attempts have been made to link this growing propensity toward crime with the breakdown of the American family. According to a 1988 survey by the National Center for Health Statistics, children in single-parent families, many of whom are products of divorce, are two to three times as likely to have emotional or behavioral problems as children in two-parent families. This paper is an attempt to solidify the correlation between divorce and juvenile crime.

ginning or end, with the remaining sentences devoted to presenting the reasons. In an essay of several pages, each paragraph would have a thesis sentence, with the remainder of the paragraph devoted to explaining and supporting that thesis. Each paragraph’s thesis sentence would represent a point in support of the paper’s overall thesis.

In an essay that has several sections or in a book with multiple chapters the organizational structure will follow this same hierarchical pattern.

The third feature of good economic writing is that it should be **solidly developed**. Once you have determined your major points, you need to explain each of those points in detail and support them with evidence. In

Figure 5.3 Structure of a Written Paper

other words, the same hierarchical structure that organizes the overall paper should also be present at the micro level. Each section of a paper should have its major point stated in a single sentence. The remainder of the section should explain the reasoning for that major point. This is illustrated in Figure 5.3, where each major point is supported by several minor points.

The remaining two features of good economic writing will be explained in the later sections of this chapter on revising and editing your paper, since they are not critical for writing the first draft.

Getting the Ideas Down on Paper

When you start writing your draft, don't worry about grammar or mechanics; don't worry about getting the details right or fixing mistakes. In fact, concerning yourself with those is a good way to get bogged down. Rather, focus on just getting your basic ideas down on paper. Indeed, it is best to do this in one sitting (or for a longer paper, one complete section at a time). McCloskey (2000) suggests that if you can't draft the paper in a

single setting, it is helpful to write yourself notes about what you have done and where you are going next, before you end your work session.

The idea is to prevent yourself from running into "writer's block." All writers experience this at some point, but if you can put this off until the revising stage you will be ahead of the game. (See Figure 5.4 for some practical suggestions for getting over writer's block.)

Many writers find it useful to create an outline of what they are trying to say. This is not a formal topic-based outline with roman numerals like you were taught in high school. Rather, Booth et al. (1995) argue that point-based outlines, which are organized as a series of points or assertions, tend to be more useful since they illustrate your argument. Note also that this step overlaps somewhat with the earlier one about developing your argument.

Sometimes it makes sense to develop the outline after you've written an initial draft of your ideas. For example, I usually find it easier to brainstorm about the first draft, writing down my ideas, and only then creating an outline to organize them. Writing is a personal process. There is no one right way to do it! We noted in Chapter 4 that the outline function of your word processing software is a helpful way to proceed regardless of the outline format you choose.

Booth et al. (1995) identify several poor schemes for organizing a paper, which you should avoid. The first is repeating the assignment: providing exactly and only what is asked for in the assignment, and in the order it was asked for. The second is merely summarizing your sources. Remember, a good paper should analyze rather than just report what others have said on a topic. The third is explaining the steps in the process that you went through to write it. A research paper should explain what you found, not what you did.

Giving Credit for Intellectual Property

All writers know to avoid plagiarism. What may not be clear is exactly what plagiarism is. Plagiarism is taking credit for someone else's words or ideas, even when it's unintentional. It is a form of academic dishonesty.

There are two types of plagiarism. The first is using someone else's words as if they were your own. That is, it means directly repeating words from another's work without including quotation marks. This occurs if someone copies directly from another's work or if he or she submit as their own a complete paper they obtained from the Internet or from some other supplier. Plagiarism can also occur if you don't keep careful records as you put together your research materials and mistake a quotation as your own paraphrase. This is plagiarism even if you reference the original source.

NOTES FOR NOVICE RESEARCHERS

A Caution About Quotations

You should use quotations cautiously, and only if you cannot possibly paraphrase without losing either meaning or impact. Excessive use of quotations suggests to the reader that you do not understand the source well enough to paraphrase and so you must rely on the words of the author. As McCloskey (2000, 45) states, "No college [or higher-level] paper can be fashioned by stringing together passages from other writers." In short, use quotations like spices, in moderation, to enhance but not overwhelm the flavor of your dish.

The second type of plagiarism is using someone's unique idea without attribution. By unique I mean something that is not common knowledge in the discipline. The key question here is what is known in the discipline. Even if something was unknown to you before your research, if it was known by experts in the field, it is considered common knowledge. You do not need to cite it to avoid plagiarism. A rule of thumb is that an idea that occurs in the publications of three different authors is considered common knowledge. Note, however, that in a literature survey you may wish to cite important studies, even if they are widely known, since showing your knowledge of important works in an area establishes your credibility. Booth et al. (1995, 167) state: "You also plagiarize when you use words so close to those in your source, that if you placed your work next to the source, you would see that you could not have written what you did without the source at your elbow." This is a strict standard, but Booth et al. provide some helpful suggestions for how to practice it.

For example, when you use someone else's work, say so early on in your discussion, rather than in a footnote at the end of the section. It is a common practice to place a footnote at the end of the first sentence of a section, using wording such as "This section draws heavily from Booth et al. (1995)."

The key to avoiding plagiarism is keeping careful track of what your sources say and what you've interpreted from those sources, and giving credit where credit is due. When you are going to cite an idea in a scholarly work, you should reference it briefly in the form of a note, then in more detail at the end of the work in the reference list or bibliography. Appendix 3A outlined the major citation styles used in academia. You should

find out from your instructor what style he or she wishes you to use, and then follow that carefully.

Revising the Paper

No one, not even a Nobel laureate, can write a perfect paper on the first try. Every first draft can be improved by revision. In fact, Booth et al. (1995, 171) observe that

Perhaps the biggest difference between experienced writers and beginners is their attitude towards th[e] first draft. The experienced writer takes it as a challenge: *I have the sketch, now comes the hard but gratifying work of discovering what I can make of it.* The beginner takes it as a triumph: *Done! I'll change that word, fix this comma, run the spell-checker, and <Print>!* A first draft is indeed a victory, but resist that easy way out.

We commonly make a distinction between a first draft and the final draft. In practice, good writers go through more than two versions of their papers, usually many more. Each major draft may be the result of several revisions or minor drafts. For example, consider the first rough draft you write versus the more polished "first draft" you're willing to have a friend read. This is analogous to the way software releases are named: WordPerfect 5 versus WordPerfect 6 are considered major revisions, while WordPerfect 6.1 versus WordPerfect 6.2 are minor revisions. In the same way, it makes sense to differentiate between major drafts of your paper, on the one hand, and revisions or versions of a major draft, on the other.

As students of economics, we should understand that the marginal benefit of revision is positive, but diminishing. The second major draft is usually much better than the first. The third major draft is even better than the second, though the improvement may not be as great. If you only write one draft, you forgo those substantial improvements. One of the most disappointing comments I write on students' final papers is "This would make an excellent first draft." In other words, it has great potential but that potential was not realized. Of course, to attain this potential you need to start drafting far enough before your deadline to have the time for multiple revisions. You know this! But here is something you may not have thought seriously about: Your instructor gives you a month to complete a writing assignment, not because she wants you to wait three weeks to start, but because a month is how long it takes to do the assignment well, by writing multiple drafts and revisions. Would you intentionally show up late to an examination and so leave half the questions unanswered because you ran out of time? Of course not, but that is effectively what you do when you wait until the last minute to begin writing the paper. Think about it.

Is the Thesis Clear?

The purpose of revision is to craft your paper so it better embodies the features of good writing listed in Figure 5.1, in order of importance. Start with focus. If the paper has no focus, then the remaining issues are irrelevant. Is the thesis or principal assertion of the paper clear? Can you underline the thesis sentence in the paper's introduction? If not, the most important thing you can do to improve the paper is fix that. Can you underline the corresponding sentence in the paper's conclusion? If not, the reader may find that your paper lacks closure, that it did not do what it set out to. Booth et al. (1995) point out that these two thesis sentences should match each other, or at least not contradict one another. If the first is framed in terms of a question, then the second should answer the question. If they do not match or if the second does not answer the question posed in the first, they need to be reworded. For example, the thesis expressed in a paper's introduction could be:

This study will attempt to determine the major factors that influence the retail demand for diamond jewelry.

The corresponding sentence in the paper's conclusion could be:

This paper has found that the demand for diamond jewelry is highly sensitive to the price, but highly insensitive to income.

Note that it is not unusual for authors to have to make this kind of correction, since when they drafted the introduction they may not have known exactly how the paper would end up. But it's important to make sure the introductory and concluding thesis statements match by the time the paper is completed.

Is the Paper Well Organized?

The next step is to review your paper's organization, which is another way to think about the paper's argument. Whether you are revising a relatively short paper, no more than a few pages in length, or a longer paper, the principles are the same. One way to check your paper's organization is to create a paragraph outline. A paragraph should be the explication of a single thought or idea or theme. Each paragraph should have a thesis sentence, typically the first sentence. If a given paragraph discusses two or more themes, you should split it into two or more paragraphs. Identify the thesis sentence of each paragraph, and copy each one into a new document in the order in which they occur. Now read the sentences and check to see if they make sense in that order and lead logically to the thesis as a conclusion. If they do not, then re-organize the order of the sentences until they do. If any sentence does not seem to contribute to the argument,

omit it. When your paragraph outline makes sense, then reorganize your paper to match the paragraph outline.

Let's try this with an example. Suppose we develop the following paragraph outline from a draft of a paper:

1. HIV/AIDS infection reduces people's ability to work when they become ill.
2. The deaths caused by HIV/AIDS reduce a nation's human capital, since not only is the labor force reduced, but the investments in those workers' human capital are also lost.
3. Immigration has been increasing into many developing nations that have high rates of HIV/AIDS.
4. A nation's labor supply depends on its birthrate, immigration rate, labor force participation rate, health expenditures, life expectancy, and investments in education.
5. The rate of HIV/AIDS infections tends to reduce labor supply.
6. The HIV/AIDS illness requires that resources be used to treat HIV/AIDS patients that could otherwise be used for disease prevention.

The principal assertion of this paper is sentence 5. How might we reorganize the sentences to support this thesis? Though there may be more than one way to do this, the following is one organizational scheme that works:

1. The rate of HIV/AIDS infections tends to reduce labor supply.
[Thesis sentence]
2. A nation's labor supply depends on its birthrate, immigration rate, labor force participation rate, health expenditures, life expectancy, and investments in education.
[Since labor supply is the key issue, what affects labor supply?]
3. HIV/AIDS infection reduces people's ability to work when they become ill.
[Reducing labor force participation]
4. The illness requires that resources be used to treat HIV/AIDS patients that could otherwise be used for disease prevention.
[Reducing the labor force participation of workers who get other diseases]
5. The deaths caused by HIV/AIDS reduce a nation's human capital since not only is the labor force reduced, but the investments in those workers' human capital are also lost.
6. [Therefore,] the rate of HIV/AIDS infections tends to reduce labor supply.
[Restatement of the thesis as a conclusion]

Note that sentence 3 in the original outline did not lead to the thesis and so was omitted.

Are Your Points Supported with Evidence?

The third step is to examine the development of each major point in your paper. Each point is itself an assertion that needs to be supported with evidence. If the first sentence in a paragraph spells out the main point of that paragraph, the remaining sentences should flesh out and support that main point. Does the main point need to be explained in more detail? Can you provide examples of what the main point says? What makes you think that the main point is valid? What evidence can you provide to bring the reader to that conclusion? Any sentences that do not contribute to this task should be omitted or moved to another paragraph where they are appropriate.

Take the second sentence in the revised paragraph outline as an example:

A nation's labor supply depends on its birthrate, immigration rate, labor force participation rate, health expenditures, life expectancy, and investments in education.

This could be fleshed out as follows:

Labor supply depends on population, which is a function of the birthrate, the immigration rate, and the death rate. The labor supply also depends on the percentage of the population that is working. This labor force participation rate depends on cultural factors, as well as expenditures on health, which help people avoid illness. Health expenditures also affect life expectancy, which influences the death rate. Though these factors determine the number of workers, the effective labor supply depends additionally on the amount of education that workers receive.

We need to address one last topic before we move on to discuss issues of style. Writer's block is a problem that afflicts all writers at some point. Figure 5.4 lists some practical tips for overcoming it, once you have a first draft to revise.

Writing Style

Once your paper has a well-defined focus and organizational structure, you can address the remaining features of good writing: clarity, conciseness, and precision. This is where you polish your argument to make it as persuasive as possible. A good writing style enables your reader to easily comprehend your writing without having to work at it.

Figure 5.4 Practical Tips for Overcoming Writer's Block

1. Copy the section of your paper where you find yourself stuck onto a separate standalone document. Edit that document, in any way that makes sense, even by radically revising it. The fact that it's a different document from the source will free you to consider radical revisions. Revise the paper until it's completely to your satisfaction. Then paste it back into the source paper, replacing the original section.
2. Print your paper (double-spaced), and review it. Changing the format from computer screen to paper helps to make ideas for improvements leap out at you.
3. Put the paper aside for a few days or a week. Let your subconscious work on it. Almost always, when you come back to it you will be able to proceed effectively.
4. Construct a paragraph outline for the paper. Summarize each paragraph using a thesis sentence. Revise the list of thesis sentences as necessary, either by adding to them or splitting them into separate ideas. Review your revised list of thesis sentences for order, redundancy, missing links, and dead ends. Use what you've learned to revise the original.

Strive for Clarity

The most important feature of good writing style is clarity. Imagine listening to a radio station that you can't quite tune in. As McCuen et al. (1993, 2) note, "You listen intensively for a short while, but you keep missing words; you can't quite understand [what the announcer is trying to say]." Unclear writing is exactly like that.

To write clearly, you should follow a few simple rules. A sentence consists of a subject and a verb. Compose each sentence so the subject is the main actor of the story, and the verb is the main action. In addition, whenever possible choose strong verbs over weak ones. You might think this would be obvious, but for some it is not. Here is an example from a research paper:

"Noticeably different is the negative coefficient of CPI."

Note how much clearer the sentence reads when the subject, *Noticeably different*, and the object, *negative coefficient of the CPI*, are reversed:

"The negative coefficient of CPI is noticeably different."

This improvement results because the main actor in the sentence is the “negative coefficient,” but originally it was shunted to the end of the sentence. This sentence also employs a weak verb: *is*. Novice writers tend to overuse *to be*, which generally only lengthens a sentence without adding anything. Verbs like *to be* tend to convey less information than stronger verbs. Suppose we replaced *is* with the stronger verb *differs*:

“The negative coefficient of CPI differs noticeably.”

Subject: “negative coefficient”

Verb: “differs”

How?: “noticeably”

Most readers would find this version much clearer than the original.

Scholarly writing has a reputation for being difficult to understand. Though it is true that scholarly writing makes use of specialized vocabulary and specialized forms of argument and evidence, it should not and need not follow the turgid academic style we see too often in scholarly works. This style has many labels, none of them complimentary. McCloskey (2000) describes it as the author posing as “The Scientist” or “The Scholar.” Lanham (1992) calls it “the Official Style.” Harvey (2000) describes this style as employing “obfuscation, nominalization, passive voice, [and] long, wordy passages which muddy up the question of who did what.” Consider, for example, this passage from a research paper:

“Even though the numbers show sustained growth the opinions of many scholars are still conflicting.”

What does this mean? I don’t know.

Researchers write this way for several reasons. First, they think they are supposed to. They think it sounds more formal or scientific than plainer prose. They think it makes them appear more knowledgeable about their subject. This is especially true of novice researchers.

Second, they may be trying to express complex ideas that they do not fully understand. McCloskey (2000) labels this the “This-Stuff-Is-So-Complex-I-Can’t-Make-It-Clear” approach to writing. The solution to this problem is straightforward, though not necessarily easy. You must think through the ideas until you understand them well enough to explain them clearly. You may need to do more research to see how other authors have explained these ideas. You may also need to discuss the ideas with colleagues or your instructor. In short, the solution is to put more work into the project.

Third, as Harvey (2000) points out, researchers may be trying to avoid taking responsibility for what they write. Writing is risky, as we noted in

Chapter 4. Writers worry that what they say might be wrong. If they write in a vague, incomprehensible style using words and phrases that they have seen in scholarly writing (e.g., *rational expectations*, *first-order conditions*, *statistical significance*), then the reader won’t be able to criticize what he or she doesn’t understand. This is like the student who, unsure how to spell a word, hedges his or her bets by spelling it several ways in an essay. In fact, the only thing that is certain is that the student will spell the word wrong. Writing without clarity is going to be criticized not for what it says, but because it says nothing. Indeed, Harvey (2000) observes that “Without clarity, you’re not really communicating, just going through the motions.”

Do not waste your efforts trying to write this way. If you focus on clarity, conciseness, and precision, your style will take care of itself.

Use the Active Voice

Academic writers often use the passive voice. Booth et al. (1995) suggest that they do this because they think it makes their writing sound more objective. The passive voice occurs when writers make the subject the recipient rather than the performer of the action (Harvey, 2002). Usually, this is done by replacing an active verb with a form of the verb *to be*. Consider the following examples:

“Next, the data were analyzed.”

“It is expected that price will increase with disposable personal income.”

“It is argued that . . .”

The solution is to rephrase the sentences using an active voice, where the subject performs the action:

“I analyzed the data.”

“We expect that price will increase with disposable personal income.”

“We argue that . . .”

Notice that the active voice is more precise in meaning. With the passive voice, the reader needs to figure out *who* is analyzing the data, *who* is expecting the price to increase, *who* is arguing. To write more precisely you must refine your thinking. For example, why do you expect that prices will increase? A likely reason is because economic theory predicts it. If so, you should explicitly say that, something our author did not. Consider the third example: who or what is arguing that . . . ? Is it the researcher? Is it the results of the analysis? Again, by thinking about subjects and verbs as you refine your thinking, you also refine your writing.

NOTES FOR NOVICE RESEARCHERS

Using the First Person

Researchers are often reluctant to use the first person in their writing. Frequently, the result is that they use the passive voice. In fact, there is nothing wrong with using the first person, especially when you are describing your interpretations and conclusions. If you can make your writing clearer by using the first person, then do so.

Passive voice is not incorrect per se, but by disguising the subject of the sentence, it makes the reader work harder to figure out what exactly you are saying. For this reason, until you become a skilled writer you should probably avoid the passive voice.

Describe Action with a Verb

Academic writers also tend to nominalize, that is, turn verbs into nouns. Again, their motivation seems to be to make the writing sound more scholarly. In fact, it merely makes it sound more pretentious and harder to read. Here are some examples:

"There *was a failure* of the results to confirm the hypothesis."

"The growth of an economy *is somewhat fluctuated* due to the amount of technological increase."

"Next, the data *were subjected* to analysis."

Nominalization tends to add words to a sentence without adding meaning. Thus, the sentence is harder rather than easier to read. Additionally, like the passive voice, nominalization obscures the subject of the sentence.

The solution to nominalization is to describe the action with a verb.

"The results *failed* to confirm the hypothesis."

"The growth of an economy *fluctuates* due to the amount of technological increase."

"Next I *analyzed* the data."

In sum, to write clearly choose strong verbs to describe the actions in your sentences, and make the main actors the subjects of your sentences.

Be Precise and Concise

Most of our discussion of style has focused on clarity, which is its most important feature of style. Now, let's discuss precision and conciseness. Dis-

cussing complex issues often requires nuance. Thoughtful writers understand that word choice matters and that synonyms have slightly different meanings. If you want readers to understand you, choose the word that means exactly what you wish to say. Precision, like clarity, suggests to the reader that you know what you are talking about. Vagueness conveys the opposite impression. Consider the following sentence:

"Some areas that enforce affirmative action are helping those who are at a disadvantage to overcome discrimination."

Think about this sentence. It begins with a vague subject: *Some areas*. One cannot easily tell whether this refers to areas of law, areas of culture, or geographic areas. The verb in the sentence (*are helping*) is weak. The next phrase (*those who are at a disadvantage*) doesn't really add any information for the reader. Compare the original with this revised version:

"The states that enforce affirmative action help individuals overcome discrimination."

Another characteristic of good writing style is conciseness. Try to make every point as concisely as you can. Avoid empty words that don't add information for the reader. Never say more than you need to.

Occam's razor is a well-known corollary to the scientific method. It proposes that when choosing among alternative theories, one should select the theory that explains the phenomenon being studied with the least complexity. The same rule is true for writing. Less is more as long as the explanation is complete.

In the real world, a supervisor will never ask you to write a report of *at least* twenty pages. Time is a scarce resource and people prefer brevity, so long as they still get the point. The same is true of academic readers.

Don't worry about the length of your paper. Never add "filler." If you go through the writing process described here, you will have enough to say to satisfy your assignment. Of course, this assumes that you have thought through your topic carefully and completely during the drafting and revision steps. If you haven't, filler won't help.

If you want your writing to be persuasive you should avoid emotion-laden phrases. Good writing is hard work. If you don't take your writing seriously, why should the reader?

Writing Mechanics

The last step in completing a paper is editing. Now is the time to worry about correcting grammar, mechanics, as well as spelling and typographical errors.

Use Complete Sentences

Serious writers employ complete sentences. A complete sentence implies a complete thought, while a sentence fragment implies fragmented or incomplete thought. As we noted previously, a complete sentence consists of a subject and a verb. A sentence fragment is a subject without a verb or a verb without a subject. Few of us would write “An economist” or “Tested the hypothesis” by themselves. Yet sentence fragments are common among novice writers.

Often novice writers will use a phrase as a sentence, when grammatically it is not. Take the following, for example:

“The key question for determining whether you have statistically significant results.”

In this phrase, *The key question* is a noun, and *for determining* sounds like a verb, but grammatically it isn’t. Compare the phrase with this sentence:

“The key question for determining whether you have statistically significant results is does your estimated t-statistic exceed your critical t.”

You can see that the original phrase, even though it has a noun and a verblike word, serves only as the subject of the full sentence. Since the original phrase includes a subject but no verb, it is not a complete sentence.

Alternatively, a phrase may include both a noun and a verb, but they are subordinated by another word: “*Although* the estimated coefficients were negative.” In this phrase, *the estimated coefficients* is a noun, and *were negative* is a verb. The phrase *the estimated coefficients were negative* is therefore a complete sentence, but by preceding them with *Although* you create a dependent clause, which is not a complete sentence because it needs additional wording to complete its meaning. Compare the original with the following sentence:

Although the estimated coefficients were negative, they were not statistically different from zero.

You should be able to see that the dependent clause was an incomplete thought, and thus an incomplete sentence.

The last type of sentence fragment occurs when a phrase has both a noun and a verb, but the verb isn’t properly conjugated for the noun: “A software package generating statistical results.” *A software package* serves as a noun, and *generating* serves as a verb, but the entire phrase isn’t a complete sentence. To determine whether you have a complete sentence ask yourself whether the phrase standing alone is a complete thought. If so, you have a complete sentence. If not, you need to fix it.

Don’t Let Sentences Run On

A run-on sentence is two or more independent clauses that are not separated with the proper punctuation. In a run-on sentence several complete thoughts are jammed together, making it difficult for the reader to determine where one ends and the next begins:

“An increase in the price of seed caused an increase in the price of corn an increase in the price of corn caused a decrease in the quantity demanded.”

A run-on can be fixed in several ways. The two clauses could be connected with a semicolon:

“An increase in the price of seed caused an increase in the price of corn; an increase in the price of corn caused a decrease in the quantity demanded.”

They could also be connected by a conjunction such as *and*:

“An increase in the price of seed caused an increase in the price of corn, and an increase in the price of corn caused a decrease in the quantity demanded.”

Alternatively, the two clauses can be divided into two complete sentences:

“An increase in the price of seed caused an increase in the price of corn. An increase in the price of corn caused a decrease in the quantity demanded.”

Given the widespread availability of spell-checking software, there is no excuse for misspelled words in a paper. McCloskey (2000, 27) points out that spelling errors and typos make you “look like a careless dolt.” At best, readers will think you are careless. At worst, they will question your entire paper. Note, however, that the spell-checking software is not sufficient to catch all typos. For example, it will not find missing words. For that you need to review the document closely. Also, it is a good idea to have someone else read the final document. A friend is more likely to catch errors than you are, since you are (too) familiar with your writing.

It is true that content is more important than mechanics in a paper. But in practice, failure to write well, that is, failure to use proper style and mechanics, inhibits your ability to communicate. This is not “mere grammar” but rhetoric and argumentation, the art of using words to persuade. If you can’t communicate, you can’t persuade the reader that your argument is correct.

SUMMARY

- Economic writing is writing that applies economic analysis to derive insights about an issue or problem.
- Writing a first draft is hard, even for experts. The key is to sit down, organize your research materials, and do it.
- Well-written papers are focused, clearly organized, fully explained, polished, precise, and concise. Additionally, they are free of grammatical and typographic errors.
- Plagiarism means taking credit, even unintentionally, for the words or ideas of others.
- No one can write a perfect paper in one draft. Good writers write many drafts, revising and revising again, until they get it right.

NOTES

1. Note that the humanities, since they are nonscience disciplines, do not use empirical evidence.

SUGGESTIONS FOR FURTHER READING

Booth et al. (1995)—Chapters 11–15 explain virtually every aspect of writing a college-level research paper, from pre-writing to revising for organization and style.

Harvey (2000)—Excellent guide to writing that is very readable, even enjoyable. Harvey includes numerous excellent examples.

McCloskey (2000)—Classic monograph on writing for economists. Readable in one sitting. Originally published in 1987 under the title *Economical Writing*.

McCloskey (1999)—Four-page executive summary of McCloskey (1987). It's better than nothing, but read the original.

McCuen et al. (1993)—Excellent guide to technical writing and speaking for engineers and other scientists.

Thomson (2001)—Useful guide for young Ph.D.s or others who are writing about theoretical economics. Quite technical, but other researchers may find it of value as well.

EXERCISES

1. Find an example of “social criticism” written by an economist (e.g., Milton Friedman, Paul Krugman, Robert Heilbroner). Find an example of a research report written by the same author. In what ways are the two types of writing similar? In what ways are they different?
2. Using the sources you have discovered so far on your research topic, draft a survey of the literature of no more than two pages in length. The focus of your survey should be to summarize the consensus in the literature on your research topic. Briefly summarize the major studies that have contributed to this consensus. Be sure that your survey illustrates the first three features of good economic writing: focus, organization, and solid development of the main points in the survey.
3. Write a clear thesis for your research paper. The thesis may well be a tentative one at this point.
4. Find a paper you have written for another class. Underline the thesis of the paper. Is the thesis clear? If not, revise it to make it clearer. Construct a paragraph outline of the paper. Evaluate the outline. Is the outline organized in the best way to lead logically to the thesis as a conclusion? If not, reorganize the outline. Are there any points that are extraneous? If so, delete them. Are all the major points in the paper explained fully and clearly? If not, correct them.
5. Print a copy of the literature survey you wrote for Exercise 2, or another short paper. Circle every usage of the passive voice with red ink. Rewrite each using active voice. Compare the passive version with the active one. Next circle every nominalization with blue ink. Rewrite each using a verb to describe the actions. Print a copy of the revised paper and compare it to the original. Which do you think is clearer? Why?

Critical Reading or How to Make Sense of Published Research

*"The difficulty you have in reading
doesn't necessarily reflect on your mental
ability. Serious reading is hard work."*

SUSAN WISE BAUER

Professional writing in any discipline can be daunting to decipher. It's kind of like your first course in a foreign language. Remember how lost you felt at the beginning? That feeling is completely understandable. Addressing this issue for students, Schroeder et al. (1985) observe, "You are typically confronted with terminology and analytical techniques with which you are at best unfamiliar, or at worst totally ignorant. There is no reason for you to feel badly about this state of affairs; as a neophyte in the world of research, you cannot expect mastery of all phases of research from the start."

Research papers tend to be written in a formal style that is designed to provide clarity and precise meaning in as concise a way as possible. This implies, however, that authors may leave out descriptions of common steps in the research process. This adds to the difficulty for novice readers, who are not (yet) members of that particular research community.

Fortunately, like a foreign language, once you learn the basic vocabulary and syntax, professional economics writing will begin to make sense. In other words, once you learn the "code," you will find it much easier to decipher published research.

Making Sense of Published Research

There are at least two ways to think about how to make sense of published research. We can call them **format** and **argument**. Let's consider them in order.

Understanding Format

What's the difference between an essay you write for an English composition class on what you did on your summer vacation and a lab report for chemistry? The answer is: specific format. The scholarly writing for each discipline uses a somewhat different format, with the physical sciences using the most structure (e.g., the lab report), the humanities using relatively less structure, and the social sciences lying somewhere in between. Even within the social sciences formats vary, though relatively less than the formats of the sciences and the humanities differ.

Economics research papers tend to follow a common format that illustrates the scientific method. This specific format allows experts to look quickly at a paper to see if they want to spend the time to read it more closely. For example, Thomson (2001, 2) states:

A reader who has found your central point interesting and wants to know more about it but has little time to invest in your work . . . should be able to grasp the novel aspects of your [research] without actually reading the paper. A lot can be learned from a well-written argument by just glancing at the way it is structured.

In economics, there are three types of scholarly works. The first is a survey of the work of others. You can think of this as an entire paper devoted to reviewing the literature on a topic; its objective is to summarize for readers what is known on a subject to date. Survey papers, unlike the other two types, are secondary literature.¹ We noted in Chapter 3 that the *Journal of Economic Literature* tends to publish survey articles. For example, the March 2002 *JEL* includes an article entitled "Looking Inside the Labor Market: A Review Article." The second type of scholarly work in economics is the purely theoretical study. It creates or modifies a theory and discusses its implications, but contains no empirical testing. An example of a purely theoretical article is Krugman's (1979) seminal article on the possibility of governments effectively practicing strategic trade policy. The third type of scholarly work in economics—and the one we will focus on—is the empirical study. The majority of research projects in economics are empirical studies. This type offers a particular advantage, especially for beginning researchers: when you do an empirical study, whether your empirical testing confirms your hypothesis or rejects it, you have results that enable you to

satisfactorily conclude your study. By contrast, when you attempt a purely theoretical project, if you can't get the theory to work the study has failed. For this reason, purely theoretical works are best left to experienced researchers.

Let's examine the format for the typical empirical research study. The typical study in economics has four essential components: an introduction, an analysis of the problem, an empirical test of the analysis, and a conclusion. An actual paper (or book) may have more formal sections, but to be credible to experts in the field the study, it must cover these four components. Let's examine each of these in turn.

The introduction should define the general topic and the specific research question, as well as explain the motivation for the research. The introduction should also include a review of the work of previous researchers on the topic, especially what is lacking in the existing literature and how the current study proposes to address that shortcoming. In a thesis or book, the literature review is usually placed in its own chapter, separate from the introduction. In a paper it is more often combined with the other introductory material.

The second section in an economic research report is the heart of any economics study: the application of economic analysis to shed light on the research question. It may be labeled "Theoretical Analysis" or "Theoretical Model." (Exactly how one applies economic analysis to a problem is the subject of Chapter 7.) This section develops the theoretical model used by the study, and derives the testable implications or hypotheses of the model.

The third section explains how the proposed analysis from section two is tested. (This will be the subject of Chapters 10 and 11.) This section could be labeled "Empirical Analysis," or "Empirical Model." It explicitly states what results would confirm the theory. It presents the results obtained from the testing procedures and interprets them. To what extent is the theory confirmed?

The concluding section in an economic research report explains the insights learned from the research. What answer did economic theory suggest for the research question? Was this answer confirmed by the empirical evidence? If not, why not?

Economics studies may also include several other parts, such as an abstract, references, and appendices, but these four components are the principal parts of the work.

Evaluating the Argument: Reading Critically

Once you realize how a research report is laid out, you are part way to understanding it. To complete that understanding, you will need to learn

how to read deeply and critically. And this requires that you understand the article's argument.

Reading critically is challenging. The purpose of reading scholarly texts is not simply to get through the pages while recording the facts the author presents. Rather, the task is the more complex one of discerning and evaluating the author's argument. Mursell (1951, 58) describes this process as follows:

When you read properly [i.e., critically], you are not merely assimilating. You are not automatically transferring into your head what your eyes pick up on the page. What you see on the page sets your mind at work, collating, criticizing, interpreting, questioning, comprehending, comparing.

To this end, scholars don't merely read texts; they study them. First, they may skim an article to see if it is going to be useful to them, paying particular attention to certain sections. For example, some scholars focus on the introduction and the conclusion; others focus on the introduction and the empirical results section. If they decide the article may be useful, they read it again more carefully, often several times. Scholars read slowly, and try to discover the meaning. They attempt to interact with the text, essentially engaging in a "conversation" with the author. This conversation takes place partly in the reader's mind, and partly through the notes the reader makes (e.g., in the margins of the text) as he or she works through the text. (How to take research notes effectively is the subject of the last section of this chapter.)

The text presents an argument. The reader responds by asking questions and expressing tentative ideas about meaning: "Is this what you mean?" "But what about this?" "Okay, I understand the previous part, but I'm not sure I fully understand this one yet."

In Chapter 4, we explained how writers develop arguments through the process of composition. Readers analyze texts in the same way, by writing ideas down, rearranging them, thinking about them, and asking questions about the tentative meaning. The meaning becomes clear after you reread an article several times, sometimes over what may be an extended period.

Bean (1996, 136) asserts that readers play two opposing roles in this process: "the open-minded believer who can succumb to the text's power and the skeptical doubter who can find weaknesses in the text." Since an argument is an assertion rather than a fact, and since, as a consequence, "every author necessarily distorts his or her subject," healthy skepticism is

NOTES FOR NOVICE RESEARCHERS

Several Important Distinctions for Critical Reading

Ruggiero (1998) suggests several important distinctions that are important for reading critically. The first distinction is between the person speaking and the idea he or she is expressing. (This is related to the ad hominem attacks we mentioned in Chapter 4.) Critical reading requires that you give a fair hearing to the ideas, even if you have a poor opinion of the author. The second distinction is between matters of taste (i.e., pure opinion) and matters of judgment. There are no grounds for criticizing the former but may be many for criticizing the latter. The third distinction is between fact and interpretation. Often, writers will present an interpretation as a fact so as to lend it more weight. Careful readers can distinguish between the two. The fourth is between literal and ironic statements. The latter are intended to provoke a strong reaction for rhetorical purposes. Fortunately, you are not likely to run into this very much in scientific writing. The last distinction is between an idea's validity and the quality of its expression. In principle, you should evaluate the validity of an idea independently of how well or poorly it is expressed. In practice, however, the quality of written or oral expression almost certainly colors your impression of the validity of one's thought. Note that this is a point raised by McCloskey (1998), as discussed in Chapter 1.

called for (Bean, 1996, 140). Careful reading of a text requires you to give adequate attention to both roles. Making sense of scholarly writing is a lot like peeling an onion: it has many layers. You shouldn't expect to discern all the levels of meaning without reading the work many times.

Questions to Guide Critical Reading

Though reading scholarly writing is rarely easy, one's ability to do it successfully does improve with practice. Critical reading is a learnable skill.

Recall that we defined an argument as an assertion supported by logical or empirical evidence. Table 6.1 lists a series of questions to guide your reading and help you identify an author's argument. Note that there may be more than one interpretation of the argument in a published work.

NOTES FOR NOVICE RESEARCHERS

Tips for Getting Through a Scholarly Journal Article

Novice researchers run into three common roadblocks that inhibit their ability to get through a scholarly journal article. These are: unfamiliar terminology or jargon, mathematical reasoning, and econometric issues and methods. Reading journal articles is difficult enough that the reader who confronts any of these roadblocks may be persuaded that he or she won't be able to understand the article. Don't let this happen to you.

When you run across terminology or concepts that you don't understand, write them down. Locke et al. (1998, 69–70) observe, "It may sound unlikely to you but we find that there are few instances when a single unfamiliar technical term brings reading to a complete halt. Just remember that in reading a technical text that is not your own, it is inevitable that there will be problems of comprehension. You have to puzzle them out or, failing that, flag them and get on with the task."

Sometimes the jargon is explained elsewhere in the paper. Alternatively, you should try to look up the terms. An obvious place to look is your research methodology or econometrics book. Another useful source is Peter B. Meyers's "Online Glossary of Research Economics" terms, which is available at <http://www.econterms.com>. If all else fails, you should not hesitate to ask your instructor for help.

Unfamiliar mathematics or econometrics should not prevent you from getting the gist of an article. Mathematical methods are used to develop a conclusion. You can think of them as "proofs." The mathematics shows the reader *how* the author arrived at the finding. But you don't need to understand the proof to grasp *what* was being proven. Almost certainly, if the paper is published someone will have checked the math to ensure that the proof works, so you don't need to worry about that. What you may not realize is that many readers, including professional researchers, bypass the mathematical details the first few times they read a paper.

These same points can be made about econometric issues. The purpose of statistical methods is to test hypotheses. Whether or not

you understand the particular techniques used, you should be able to determine if the results confirm or refute the researcher's hypothesis. Locke et al. (1998, 70) point out, "In any sound report, somewhere will be found a plain language description of anything found in the analysis that really mattered. . . . In many cases, that bit of text should allow you to proceed with intelligent reading—even if not with full appreciation of the elegance (or appropriateness) of the researcher's statistical analysis."

Table 6.1 Identifying the Author's Argument

1. What question is the author asking?
2. What answer does the author propose (i.e., what is the principal assertion of the study)?
3. In what ways does the study improve upon previous research?
4. How does the proposed answer compare with that provided by previous research?
5. What are the major logical or theoretical reasons for the author's argument?
6. What empirical evidence does the author provide?
7. What assumptions is the author making in his or her reasoning?

Let's work through a sample scholarly article from a professional economics journal. The article is David Romer's "Do Students Go to Class? Should They?" (1993), which is reproduced on the following pages. Read the article through at least once.

Now, reread the article carefully. As you read, think about the questions in Table 6.1 and write down tentative answers as they come to you. (The answers that I provide for each question are exact quotations from the article to demonstrate where I found the answers. When you write your own answers to these questions, you should put them in your own words. For the reasoning behind this, see the section "Taking Research Notes and Writing Abstracts and Critical Reviews" later in this chapter.)

Do Students Go to Class? Should They?

David Romer

Lectures and other class meetings are a primary means of instruction in almost all undergraduate courses. Yet almost everyone who has taught an undergraduate course has probably noticed that attendance at these meetings is far from perfect. There is surprisingly little systematic evidence, however, about attendance and its effects. There are three natural questions. What is the extent of absenteeism? How much, if at all, does absenteeism affect learning? Should anything be done about absenteeism?

This article presents quantitative evidence on the first two of these questions, and speculative comments on the third. First, attendance counts in economics courses at three relatively elite universities indicate that absenteeism is rampant: usually about one-third of students are not at class. Second, regression estimates of the relation between attendance and performance in one large lecture course suggest that attendance may substantially affect learning: considering only students who do all of the problem sets and controlling for prior grade point average, the difference in performance between a student who attends regularly and one who attends sporadically is about a full letter grade. In light of these results, steps to increase attendance, including making attendance mandatory, may deserve serious consideration.¹

1 I have been unable to find any previous investigations of the extent of absenteeism. There have been a few other studies of the relation between attendance and performance (for example, Schmidt, 1983; Park and Kerr, 1990). These studies generally confirm the findings here that attendance and performance are related even when a variety of student characteristics are controlled for. The present study differs from this earlier work in focusing on the quantitative magnitude of the relationship and on the issue of the extent to which the relationship reflects a genuine effect of attendance.

■ David Romer is Professor of Economics, University of California, Berkeley, California.

Journal of Economic Perspectives

Do Students Attend Class?

Counts were made of the number of students attending one meeting of every undergraduate economics class during a "typical" week of the spring 1992 semester at three schools. School A is a medium-sized (6000 undergraduates) private university; School B is a large (20,000 undergraduates) public university; and School C is a small (2500 undergraduates) liberal arts college. The schools are intended to be representative of the upper echelons of American colleges and universities. All three are classified by *Barron's Profiles of American Colleges* (1991 edition) as "highly competitive," the second highest of their six categories.²

The attendance counts were made a few weeks before the end of the semester at each school. This choice avoided both times when attendance is generally thought to be unusually low (such as just after exams and immediately before and after vacations) and times when it is generally thought to be unusually high (such as just before exams). Individuals at all three schools independently suggested that attendance a few weeks before the end of the semester was likely to be representative of average attendance. Attendance was taken at one meeting of each class during the sample week. Current enrollment figures were obtained from departmental offices.

Table 1 reports the results. The first row shows the overall absenteeism rates at each school. At School A, 34 percent of students were absent; at School B, 40 percent; and

2 A total of 127 schools, enrolling about 675,000 undergraduates, are classified by *Barron's* as "most competitive" or "highly competitive." Of these, 24 schools, with 140,000 undergraduates, are private universities with between 4000 and 10,000 undergraduates; 14, with 270,000 undergraduates, are public universities with over 10,000 undergraduates; and 56, with 100,000 undergraduates, are colleges or universities with minimal graduate programs and fewer than 3500 undergraduates. The remaining schools are small and medium-sized public colleges and universities (15 schools, with 70,000 undergraduates), large private universities (5 schools with 70,000 undergraduates), and small private universities (13 schools with 30,000 undergraduates).

*Journal of Economic Perspectives***Table 1** Absenteeism Rates in Economics Classes

	<i>School A</i>	<i>School B</i>	<i>School C</i>
All Economics Courses	34.0%	39.7%	24.8%
By Size of Course:			
Small (bottom 33%)	27.0	37.7	21.5
Large (top 33%)	38.8	42.9	30.4
By Mathematical Content:			
Mathematical	10.0	17.6	16.7
Non-Mathematical	34.3	41.5	25.5
By Type of Course:			
Principles & Intermediate Theory	37.3	40.5	29.7
Upper Level, Only Principles Required	26.4	41.1	17.7
Upper Level, Additional Requirements	33.3	35.2	20.7

at School C, 25 percent.³ In short, on a typical day at a typical elite American university, roughly one-third of the students in economics courses are not attending class.

The remaining rows of the table break down the overall figures along various dimensions. Course size appears to have an important effect on absenteeism. At all three schools, absenteeism is considerably lower in the smallest third of classes than in the largest third. In addition, average class size is lowest at School C and highest at School B, which is consistent with the fact that absenteeism is lowest at C and highest at B. Absenteeism is also lower in courses with a significant mathematical component (such as econometrics, honors sections of intermediate theory, and field courses in theory). This pattern holds at all three

³ Attendance counts were inadvertently not made in a handful of classes at School C. These classes do not appear to differ in any systematic way from the classes at which attendance was taken.

Journal of Economic Perspectives

schools.⁴ Similarly, at all three schools absenteeism is somewhat higher in core courses than in field courses.

Finally, it is generally perceived, not surprisingly, that students attend class more often when the quality of instruction is higher. At School B, for example, absenteeism is 34 percent for courses taught by regular faculty and 47 percent for courses with other instructors. To investigate this issue more systematically, course evaluation data for all undergraduate economics courses for one term were obtained from a fourth school, School D. This school, like School B, is a large public university. The two variables of interest are students' average rating of the overall effectiveness of the instructor and the fraction of the students enrolled in the course who returned the course evaluation form (which is a reasonably good measure of attendance at one of the last class meetings of the term). The point estimates from a simple regression of the fraction of students attending the class on the average rating imply that raising the average rating from the 25th percentile to the 75th lowers absenteeism by 10 percentage points; the *t*-statistic on the rating variable is 3.4. Thus the quality of instruction (or at least students' perception of that quality) appears to have an important impact on attendance.

Other features of the data from School D generally confirm the findings for the other schools.⁵ Absenteeism is high (45 percent across all courses), and lower in small courses than in large (31 percent in the smallest third of courses and 54 percent in the largest third). Again, absenteeism is lower

⁴ The figures for mathematical courses at School A are based on only one course. Thus this figure should be given little weight.

⁵ The data from School D are not strictly comparable with those from the other schools, because they reflect class meetings at the end of the term and because a few students are present but do not return the evaluation form. It seems unlikely that these differences have any substantial effect on the results.

Journal of Economic Perspectives

in courses with a mathematical emphasis (39 percent, versus 47 percent for other courses), and higher in core courses (52 percent, versus 31 percent in field courses that only require principles and 37 percent in advanced field courses).

A straightforward regression confirms these patterns of differences in absenteeism across different types of courses. Specifically, using the data from all four schools, I ran a regression (across courses) with the fraction of students absent as the dependent variable, and a constant, the log of enrollment, and dummies for mathematical content, for the two types of upper level courses, and for three of the four schools as independent variables. The resulting estimates imply that a doubling of enrollment is associated with a rise in absenteeism of 4 percentage points; that mathematical content is associated with a fall in absenteeism of 3 percentage points; and that moving from a core course to either type of field course is associated with a fall in absenteeism of 5 to 7 percentage points. The coefficient on the enrollment variable is highly statistically significant; those on the field course dummies are marginally so; and those on the dummy for mathematical content and the three school dummies are insignificant.

Should Students Attend Class?

These findings raise the question of whether absenteeism has a substantial effect on learning. It is possible that students do not attend class because they would learn relatively little if they did—because the instruction is of low quality, or because they have already mastered the material, or because they can learn the material better by spending the same time studying in other ways. Alternatively, it is possible that learning is severely adversely affected by absences, but that many students are absent anyway—because they have genuinely better uses of their time, or because they mistakenly believe that attendance is not important to learning, or because they attach relatively little importance to learning.

Because student attendance is not exogenous—students *choose* whether to attend class—it is not possible to isolate definitively the impact of attendance on learning.

Journal of Economic Perspectives

But this section presents some suggestive evidence. In the fall 1990 semester, I took attendance at six meetings of my large intermediate macroeconomics course. The resulting data can be used to investigate the relation between attendance and performance.

As in other courses, overall absenteeism was high (25 percent). Twelve percent of the students missed four or more of the meetings where attendance was taken; 28 percent missed two or three; and 59 percent missed none or one. Thus, absenteeism appears to be a mixture of some students missing most classes and many students missing a smaller number of classes.

Student performance is measured as the overall score on the three exams in the course. For ease of interpretation, the scores are converted to the usual 4-point grading scale: 3.84 and above represents an A; 3.50 to 3.83 an A–; and so on down to 1.50 to 1.83 for a C–. Because no D+’s or D–’s were assigned, 1.17–1.49 represents a D and 1.16 and below an F.

The first column of Table 2 reports the results of a simple regression of performance on the fraction of lectures attended.⁶ The regression reveals a statistically significant and quantitatively large relation between attendance and performance. The *t*-statistic on attendance is 6.2; the point estimates imply that a student who attends only a quarter of the lectures on average earns a 1.79 (C–), while a student who attends all of the lectures on average earns a 3.44 (B+). Attendance alone accounts for 31 percent of the variance in performance.

6 There is one econometric complication worth mentioning: because attendance was not taken at every class meeting, some of the variation across students in measured attendance is due to measurement error rather than to true differences in attendance over the whole semester. If the class meetings at which attendance was taken were a random sample of all the meetings—which appears to be a good approximation—it is straightforward to estimate the size of the measurement error. This procedure implies that 38 percent of the variation in measured attendance represents measurement error. This estimate can be used to correct the regression coefficients, standard errors, and R^2 's for the bias that would otherwise be introduced by the measurement error. All of the results reported in Table 2 have been corrected in this way.

*Journal of Economic Perspectives***Table 2** The Relationship Between Attendance and Performance

	(1)	(2)	(3)	(4)	(5)
<i>Sample:</i>	<i>Full</i>	<i>Restricted</i>	<i>Full</i>	<i>Full</i>	<i>Restricted</i>
Constant	1.25 (0.27)	1.02 (0.58)	1.07 (0.23)	-0.67 (0.32)	-0.78 (0.43)
Fraction of Lectures Attended	2.19 (0.35)	2.47 (0.70)	1.74 (0.46)	1.52 (0.32)	1.38 (0.58)
Fraction of Problem Sets Completed			0.60 (0.32)		
Prior GPA				0.78 (0.12)	0.86 (0.14)
Sample Size	195	116	195	195	116
R^2	0.31	0.26	0.33	0.47	0.48

Standard errors are in parentheses. The restricted sample consists of the students who completed all of the problem sets.

Students who are more interested in the material, or more skilled academically, or more focused on academics are almost certain to attend class more often than students who are less interested, less skilled, or less focused (other factors held constant). If this is the case, then the results in Column 1 of Table 2 to some extent reflect a general impact of motivation on performance rather than a true effect of attendance.

I attempt to address this problem in three ways. First, I restrict the sample to the 60 percent of the students who did all nine problem sets. It seems likely that most of the students who were not devoting serious effort to the course did not complete all of the problem sets. In addition, the lowest problem set score was dropped in computing the course grade; thus the students who completed all nine may have been especially motivated. On both grounds, this restricted sample may be more homogeneous in terms of general motivation than the full class. But as the second column of Table 2 shows, the relation between attendance

Journal of Economic Perspectives

and performance in this sample is actually slightly stronger than for the class as a whole.

Second, doing the problem sets is arguably as good a proxy as attending the lectures for motivation. But Column 3 shows that there is a much stronger relation between attendance and performance than between doing the problem sets and performance: when both variables are entered in the regression, the coefficient on the fraction of lectures attended is almost three times as large as the coefficient on the fraction of problem sets completed. Thus, either attendance is a much better proxy than completing the problem sets for motivation, or attendance has a large additional impact on performance.

Third, data were obtained on students' grade point averages as of the beginning of the semester. Including GPA as a control variable in the regression serves to control for some of the differences across students in general ability and motivation. In fact, because students' academic performance in previous classes depends in part on their attendance in those classes, the coefficient on prior GPA will capture some of the effect of attendance on performance; as a result, including GPA as a control variable could cause the coefficient on attendance to *understate* the true impact of attendance on performance.

Column 4 of Table 2 shows the effects of including grade point average in the regression. Prior GPA has an extremely strong relation with performance. But the inclusion of GPA has little impact on the relation between attendance and performance. The coefficient on attendance is two-thirds as large as it is in the basic regression in Column 1, and it remains highly significant. The point estimates imply that a student with the mean prior GPA earns on average a 2.13 (C) if he or she attends a quarter of the lectures but a 3.27 (B+) if his or her attendance is perfect.

Finally, Column 5 shows the results of both restricting the sample to the students who did all nine problem sets and controlling for prior GPA. Even in this case, the

relation between attendance and performance remains large and significant. The estimates imply that a student with the mean prior GPA earns on average a C+ if he or she attends only a quarter of the classes, compared to a B+ if attendance is perfect.

None of these ways of attempting to address the problem that attendance is not exogenous is definitive. Nonetheless, they all give similar results: simple ways of controlling for motivation and other omitted factors have only a moderate impact on the relationship between attendance and performance. Thus, although the possibility that the relationship reflects the impact of omitted factors rather than a true effect cannot be ruled out, it seems likely that an important part of the relationship reflects a genuine effect of attendance.

Should Attendance Be Mandatory?

Absenteeism is rampant in undergraduate economics courses at major American universities. In addition, there is a very strong statistical relationship between absenteeism and performance, and the evidence is consistent with the view that this relationship has an important causal component.

These results raise the question of whether measures should be taken to combat absenteeism. At the very least, exhortations to attend class seem called for, and those exhortations can be backed up with data. But stronger measures might be preferable. A generation ago, both in principle and in practice, attendance at class was not optional. Today, often in principle and almost always in practice, it is. Perhaps a return to the old system would make a large difference to learning. There is no way to find out but to try. I believe that the results here both about the extent of absenteeism and its relation to performance are suggestive enough to warrant experimenting with making class attendance mandatory in some undergraduate lecture courses.

One could also use mandatory attendance to perform a genuine controlled experiment that could isolate the true

impact of attendance on mastery of the material. Specifically, one could randomly divide the students in a course into two groups, an experimental group whose grading was based in part on attendance and a control group whose grading was not. By comparing the attendance and the performance of the two groups, one could learn both the impact of mandatory attendance on absenteeism and the impact of attendance on performance.⁷ Unless either the impact of mandatory attendance on absenteeism or the size of the class were very large, the results of carrying out this experiment for a single class would not allow one to estimate the impact of attendance on performance with much precision. But the pooled results from several such experiments could.

■ I am grateful to Caroline Fohlin, Matthew Jones, and Costas Tsatsaronis for excellent research assistance, and to Robert Cox, Roger Farmer, Steven Fazzari, Alan Krueger, Christina Romer, Paul Ruud, Joseph Stiglitz, Timothy Taylor, and Robert Turner for useful comments.

References

- | | |
|---|--|
| Park, Kang H., and Peter M. Kerr, "Determinants of Academic Performance: A Multinomial Logit Approach," <i>Journal of Economic Education</i> , Spring 1990, 21, 101–11. | Schmidt, Robert M., "Who Maximizes What? A Study in Student Time Allocation," <i>American Economic Review</i> , May 1983, 73:2, 23–28. |
| ■ From David Romer, "Do Students Go to Class? Should They?" <i>Journal of Economic Perspectives</i> , 7 (Summer), pp. 167–174. Reprinted with permission. | |

7. Such an experiment would presumably require appropriate approval. Students could be given the right to opt out of the experiment by being allowed to choose (before the class is divided into the experimental and control groups) to have their grade based on a formula that gave attendance half the weight used in the grading formula for the experimental group. Fairness could be ensured by assigning grades to all students using all three procedures (experimental, control, and opting out), and making the mean grade for the full class the same under all three procedures, before the allocation of the students to the three groups was known to the person assigning grades.

As you go through the Romer article, ask yourself what role each part of the text plays in the author's argument. For example, what is the author's principal assertion? What are the major pieces of evidence he uses to support the assertion? Note that, just as in many economics papers, the sections of Romer's paper are not labeled "Introduction," "Theory," "Empirical Testing," and "Conclusions," but those pieces are in the paper nonetheless.

1. **What question is the author asking?** This may be the most important question in Table 6.1. If you can't determine this, you are probably wasting your time reading the paper. The research question should be found in the introductory section of the paper. Romer's paper is a complex argument consisting of three assertions, each of which leads logically to the next. He summarizes his argument in the opening paragraph, where the key question is:

"How much, if at all, does absenteeism affect learning?"

2. **What answer does the author propose?** The author's proposed answer, that is, the principal assertion of his or her argument, can often be found in the introduction to the paper. Romer states this in the second paragraph:

"Regression estimates of the relation between attendance and performance in one large lecture class suggest that attendance may substantially affect learning."

If the author's proposed answer isn't in the introduction, it should always be clearly stated in the concluding section of the paper. In the final section, Romer recapitulates:

"[T]here is a very strong statistical relationship between absenteeism and performance, and the evidence is consistent with the view that this relationship has an important causal component."

3. **In what ways does the study improve on previous research?** The introduction to the paper should indicate how the current research is an improvement over its predecessors. In this case, the author indicates on page 1 (in a footnote):

"The present study differs from . . . earlier work in focusing on the quantitative magnitude of the relationship and on the issue of the extent to which the relationship reflects a genuine effect of attendance."

Often, though not in this case, the distinctive features of the present research are also noted in the concluding section of the paper.

4. **How does the proposed answer compare with that provided by previous research?** The answers provided by previous researchers are usually found in the literature survey in the paper's introduction, which in this case is limited to a footnote on page 1:

"These [previous] studies generally confirm the findings here that attendance and performance are related even when a variety of student characteristics are controlled for."

Additionally, the conclusion often indicates how the present results were either comparable to or different from those of previous researchers.

5. **What are the major logical or theoretical reasons for the author's argument?** The author's argument consists of the principal assertion (i.e., the author's proposed answer to the research question), along with logical and theoretical reasoning, empirical evidence, and underlying assumptions. The theoretical reasoning can be found in the paper's analytical section. Romer's theoretical argument, since it deals with a topic we are all familiar with, is largely implicit:

"Lectures and other class meetings are a primary means of instruction in almost all undergraduate courses."

This is the first sentence of the paper. It implies that class attendance leads to learning.

6. **What empirical evidence does the author provide?** Empirical evidence is found in the section of the paper on empirical testing. This includes the overall testing methodology, the results, and their interpretation. The first part of Romer's empirical argument is to estimate the extent to which absenteeism occurs:

"In short, on a typical day at a typical elite American university, roughly one-third of the students in economics courses are not attending class."

Romer then goes on to expand on this point:

"Course size appears to have an important effect on absenteeism. . . . [A]bsenteeism is considerably lower in the smallest third of classes than in the largest third. . . . Absenteeism is also lower in courses with a significant mathematical component. . . . Similarly, . . . absenteeism is somewhat higher in core courses than in field courses. Finally, . . . students attend classes more often when the [perceived] quality of instruction is higher."

In the second part of Romer's empirical argument, he reports on a series of regressions to estimate the relationship between attendance

and performance in class when controlling for a number of outside variables, including students' prior GPA and student motivation.

"A simple regression of performance on the fraction of lectures attended . . . reveals a statistically significant and quantitatively large relation between attendance and performance. . . . Simple ways of controlling for motivation and other omitted factors have only a moderate impact on the relationship between attendance and performance."

7. What assumptions is the author making in his reasoning? Romer, like all researchers, makes a number of assumptions in his argument. These include that GPA is a good measure of student learning, that the results he observed at the "elite" universities he studied are generalizable to all institutions of higher education, and that economics students and courses are similar to other college students and courses.

When you find it difficult to dissect a scholarly article, you may take heart from the following observation by Trelogan (2001):

Why did it take so long to identify the arguments and to get clear about their premises and conclusions? The answer is that there is no mechanical procedure for doing this and that quite in general, reading—really reading—is a difficult art. We had to read carefully and analytically, and we had to do some serious thinking about the content and the structure of the passage we were examining.

An additional complication, as a number of commentators have noted, is that many scholarly works are not well written.

Evaluating Published Research

Once you've identified the argument in a published work, the next step is to evaluate the argument, that is, to assess its validity and reliability. This is not a trivial task, so let's begin with the easy questions.

Does the author have an apparent conflict of interest? For example, is the research funded, sponsored, or published by an organization that might be considered less than objective about its findings (e.g., research funded by the tobacco companies that finds no adverse health effects from smoking)? A somewhat milder version of this concern: does the author or the journal promote a particular point of view (e.g., liberal views, conservative views)? If so, you should factor this into your evaluation of the work.

Is the study published in a refereed journal? We noted in Chapter 3 that scholarly articles published in professional journals are likely to have un-

dergone a formal evaluation process. Multiple experts in the field review the articles blindly, that is, without being aware of the identities of the authors so as to keep their evaluations objective.² The more prestigious the journal, the more rigorous the evaluation, and the more likely the author's argument is to be valid. Thus, articles published in the *American Economic Review* are likely to be of higher quality and greater importance than articles published in, for example, a regional journal. A similar review process takes place for academic books before they are published. Thus, when a study is published, novice researchers can assume that it has a certain amount of validity.

Another thing to remember is that the usefulness of a published study depends on more than the ostensible quality of the journal in which it is published. For example, James Tobin (1978) published an important, very widely cited article ("A Proposal for Monetary Reform") in the *Eastern Economic Journal*, a relatively minor journal. By contrast, numerous articles have been published in top journals and were never cited again, or, if they were, they haven't stood the test of time. In short, the ultimate validity of a piece of research comes only after time, when many other researchers have tested that theory in numerous different ways. Thus, when a scholar reads a research study, especially a recent one, he or she will need to critically evaluate it for him- or herself.

More Questions to Guide Critical Reading

Table 6.2 lists a series of questions to help you evaluate an author's argument. Naturally, evaluating a scholarly argument is inherently more difficult than simply identifying it in a published study. Several of the issues discussed here will be dealt with in much greater detail in the remaining chapters of this book. Here, we just introduce them.

Table 6.2 Evaluating the Author's Argument

1. Does the theoretical analysis make sense?
2. Are the data used adequate to the task?
3. Does the empirical methodology adequately test the hypothesis?
4. Are the assumptions reasonable?
5. Is the analysis (theoretical and empirical) clearly explained?
6. Do the conclusions follow from the evidence presented?
7. On balance, is the author's argument convincing to you?

As we discuss these questions, let's use them to assess Romer's (1993) article.

1. **Does the theoretical analysis make sense?** Do the reasons adequately lead to and support the hypothesis? Is the argument deep enough? Is the argument broad enough to be convincing? Are there other reasons that need to be brought into the analysis? Is there an alternative explanation of the research problem that makes as much sense as that proposed by the author?

We identified Romer's argument in the answers to questions 2, 5, 6, and 7 in Table 6.1. Paraphrasing those answers gives us:

"Lecture remains the primary pedagogical tool in undergraduate economics. Estimates of attendance in undergraduate economics courses at three top universities indicate that on average one-third of students are absent from class. Regression results on the effects of attendance on GPA, when controlling for motivation and student quality, indicate that attendance has a large and significant positive effect on student performance."

This reasoning makes sense, since each point leads to the next and ultimately to the conclusion. The theoretical foundation—that class attendance should be important to learning—is logical. (Note, however, that valid reasoning doesn't rule out the possibility that there's a contrary logical argument—for example, that students could use the textbook to substitute for class attendance.)

2. **Are the data used adequate to the task?** Are the data completely documented? Is the data set from an authoritative source? Are the data a representative sample or a special case? If the latter, are the results applicable to the situation you care about (e.g., does a study using Virginia data generalize to the United States as a whole)?

The data on class attendance were taken during one meeting of a "typical" week in all undergraduate economics courses at three "relatively elite" institutions. Though he did not identify the specific schools, Romer described one as a "medium sized private university," the second as a "large public university," and the last as a "small liberal arts college." More detailed data were obtained from another "large public university." Romer also pointed out several small shortcomings in the data.

The data used to assess the relationship between class attendance and performance were obtained from Romer's intermediate macro course. The data on attendance were taken from a sample of six class meetings.

The data were adequately documented, but it is reasonable to question how representative they might be for undergraduate economics students at schools nationwide. On the other hand, it is important to recognize that data are never perfect.

3. **Does the empirical methodology adequately test the hypothesis?** If you obtain the best possible results, how confident are you that the hypothesis is correct? Does the methodology clearly discriminate between the hypothesis and alternative interpretations of the evidence?

Subject to the possible shortcomings of the data just described, the regression analysis is a reasonable and commonly used methodology for testing questions such as the ones Romer posed. He attempts to control for motivation and student quality, which is an improvement over previous work. It is to his credit that Romer points out that his approach does not definitively deal with the problem that class attendance is not exogenous.

4. **Are the assumptions reasonable?** If the assumptions are factual, are they true? If they are empirical (e.g., perfect competition) how critical are they to reaching the author's conclusion? If they are critical, are they reasonable? For example, are they commonly made in this literature? Do the data satisfy the assumptions of the testing methodology (e.g., if the author uses an ordinary-least-squares (OLS) regression, are the requirements for OLS satisfied)?

Most researchers would accept that GPA is a reasonable, if imperfect, measure of student learning. Romer's sample of data is a special case, and thoughtful researchers will be concerned about how representative it is. However, lacking evidence to the contrary, most would accept it provisionally.

5. **Is the analysis clearly explained?** Does the description omit any steps that are necessary to reach the conclusion? (Recall that this may be a result of space limitations or instructions from the editors about omitting generally accepted warrants.) What is it that isn't clear? Is it terminology? Are all tables and graphs discussed and explained? Ruggiero (1998, 65) indicates that well-written books or articles state their principal argument clearly and visibly. Then he observes, "Not every article or book, unfortunately, is well written."

Romer's article is well written, and clearly organized, even though it does not use the typical format for a scholarly article. The main points in his argument are clearly explained; the tables serve to illustrate the argument.

6. **Do the conclusions follow from the evidence presented?** Does the author correctly interpret the empirical results? How well does the empirical evidence support the conclusion? Does the author address conflicting evidence or views? Is there an alternative explanation that fits the evidence? How do the results of this study compare with previous ones? If they are comparable, that is a point in favor of this study. If they are not, are convincing reasons supplied for why this is so? Such reasons might include that the current study has better experimental controls, a different data set, or more complete or otherwise better data.

This paper is more empirical than theoretical. The regression results strongly support the assertions that a significant portion of students do not attend class regularly and that class absences adversely affect student performance. The conclusion that faculty should seriously consider stronger measures to encourage class attendance follows directly from this evidence. Additionally, the findings are consistent with earlier studies.

7. **On balance, is the author's argument convincing to you?** In light of the reasons and evidence provided, is the argument persuasive? In other words, is the supporting evidence consistent with the hypothesis? If so, the argument may be considered valid. Does the study actually answer the research question(s)? For example, was there anything you expected to see, but didn't? Did the paper show signs of carelessness, such as failing to provide full references to all citations? Are there significant studies on this topic that the author does not appear unfamiliar with? Did the author identify limitations of the study?

Romer's argument is very well done. We can conclude that the article is convincing, subject to the limits of its methodology.

Though evaluating scholarly work requires that the reader make a judgment about its quality, Ruggiero (1998, 67) points out that it is not always necessary to agree or disagree in total with an author's argument:

If you agree in part and disagree in part, explain exactly what your position is and support it carefully. . . . If some vagueness or ambiguity in the author's argument prevents you from giving a flat answer, don't attempt one. Rather, say, "it depends," and go on to explain. . . . If you must deal with conflicting testimony and cannot decide your position with certainty, identify the conflict and explain why you cannot be certain.

Evaluating published research is as much an art as a science, and we will continue to discuss it in the remainder of this book. As Locke et al. (1998,

71) clearly state, "To be honest, it takes years of experience to quickly discern flaws of logic and imperfections of analysis in a complicated investigation." Undergraduates are sometimes reluctant to commit themselves definitively, especially when they feel they don't fully understand something. You shouldn't use either Ruggiero's or Locke et al.'s points as an excuse to avoid drawing a conclusion about a published work.

Taking Research Notes and Writing Abstracts and Critical Reviews

Earlier in this chapter, we briefly discussed the note-taking process. Let's consider it in more detail now. When you take notes on a reading you read it more carefully. The first time you read an article, you should focus on the big picture. Instead of taking notes on each paragraph, force yourself to read the entire section, then ask yourself: what did it say? Only after you can articulate the point of the section should you go back and take notes on the details.

This suggests a related point. Good readers vary their speed and depth of reading to match their purpose—whether they are interested in getting the main ideas or all the details. For example, as we noted previously, you don't need to grasp all the nuances of the particular statistical technique an author uses to get a basic understanding of the results.

The notes you take while reading a scholarly article are really addressed to you. After you finish the article, it is a good idea to transcribe your notes onto paper (e.g., index cards) or a computer file. Always start by recording the *complete* bibliographic information using the appropriate citation style. You will be amazed how much harder it is to track down an article to obtain citation information the night before a paper is due. Save yourself the trouble and get it right the first time.

Indeed, it is good practice to make copies of all articles (or chapters of books) that are important to your research project. Though this costs a bit for duplicating or printing fees, it is a small price to pay to avoid time and aggravation later. This is especially true for those articles you need to read more than once. When you copy an article, make sure you include the whole article, including references and appendices. You should also include a copy of the title page of the journal, indicating journal title, volume and issue number, and date.

As a general rule, it is better to paraphrase what you read than to record it verbatim. Notes that are collections of direct quotations suggest that you don't understand what you read or you haven't made the effort to understand. Paraphrasing requires you to process the information, which means that you will learn it better. In addition, as you paraphrase you prepare the

ideas for inclusion in your own writing. If you use a highlighter when you read, and you find yourself highlighting a great deal of the text, it might be useful to make an effort to think more about what you're reading and use the highlighter less—for example, only on the truly important aspects of the text. Finally, if you need help getting through an article, talk to your instructor.

You can make the notes you write on a scholarly work more formal by turning them into an **abstract** or a **critical review**. Both are natural products of the critical reading process. Most people think of an abstract as a summary of a scholarly article or book, but a better way to think of it is as a summary of the author's *argument* (together with complete citation information). As such, as Cohen & Spencer (1993, 223) point out, an abstract may present the material in a different order than that of the original article. Scholarly works often include an abstract, but it is good practice to write your own. If you can do it well, you clearly understand the article or book.

Let's write an abstract for Romer (1993), the article we read critically earlier. Every abstract begins with complete bibliographic information for the work. You should consult your instructor (if your paper is an assignment for a course) or the style guide (if it is a submission to a journal) to determine the correct citation style. The style used in Table 6.3 is Style B of the *Chicago Manual of Style*, Fourteenth Edition.

Recall from Chapter 4 that an argument is a thesis supported by major theoretical and empirical reasons, which are themselves fleshed out by minor reasons and other details. In an abstract of one hundred to two hundred words, you only have enough space to state the thesis and the major reasons. Based on the answers to the questions in Table 6.1, it should be easy to identify these. The thesis of the article is the answer to question 2 in the table: What is the author's proposed answer to the research question? The theoretical reasoning is the answer to question 5. The empirical evidence is the answer to question 6. Putting this all together and condensing it yields the sample abstract in Table 6.3.³

A critical review is an abstract augmented with a critical evaluation of the work. Thus, it includes complete citation information, a summary of the author's argument, and an assessment of that argument. Such an evaluation can be easily written by summarizing the answers to the questions in Table 6.2, as shown in Table 6.4.

A related research product is called an **annotated bibliography**. This is a list of references that includes a few sentences summarizing and critiquing each item. In other words, it is something like a collection of critical reviews on a single research topic. An annotated bibliography is a good

Table 6.3 Sample Abstract

Romer, David. 1993. Do students go to class? Should they? *Journal of Economic Perspectives* 7 (Summer): 167–174.

Estimates of attendance in undergraduate economics courses at three top universities indicate that on average one-third of students are absent from class. Regression results from a large intermediate macro course of the effects of attendance on GPA, when controlling for motivation and student quality, indicate that attendance has a large and significant positive effect on student performance. Romer argues, as a consequence, that faculty should seriously consider ways to increase student attendance.

Table 6.4 Sample Critical Review

Romer, David. 1993. Do students go to class? Should they? *Journal of Economic Perspectives* 7 (Summer): 167–74.

Estimates of attendance in undergraduate economics courses at three top universities indicate that on average one-third of students are absent from class. Regression results from a large intermediate macro course of the effects of attendance on GPA, when controlling for motivation and student quality, indicate that attendance has a large and significant positive effect on student performance. Romer argues, as a consequence, that faculty should seriously consider ways to increase student attendance.

The theoretical reasoning in this paper makes sense. The empirical methodology seems appropriate, and the results are consistent with the hypothesis. The principal shortcoming is the possibly nonrepresentative data sample. On balance, however, the article is convincing, subject to the limits of its methodology. Any researcher doing work in this area should be familiar with this article.

way to determine how complete your literature review is. It is also very useful for preparing the written literature survey for a research report. The Suggestions for Further Reading at the end of each chapter in this book are examples of annotated bibliographies. See also the annotated list of articles in Appendix 6A.

The only way to learn how to read and understand scholarly research is to practice this important skill. Reading through this chapter is not

enough! The more you practice, the sooner you will become adept at reading critically. To that end, in addition to the paper reproduced later in this chapter, another paper is also included on the website for your review. Please be sure to read it as well.

SUMMARY

- Scholarly literature is challenging to read.
- Scholarly writing follows a specific format like a lab report.
- You must read critically to discern and evaluate an author's argument.
- Critical reading requires judgment on the reader's part.
- A research abstract is a summary of an author's argument.
- A critical review is an evaluation of an author's argument.

NOTES

1. There is another type of scholarly work, called a meta-analysis, that is similar to a survey paper but includes original research. A meta-analysis uses sophisticated statistical techniques to analyze the separate results of previous studies. It does this to discern insights that may not be apparent from any individual study, similar to the way a mathematical average provides a single measure that summarizes a broad sample of data. Meta-analyses are much more than mere averages, however, and lie outside the scope of this book.
2. Note, however, that not all journals have a peer-review process. If an article is peer-reviewed, it is probably of high quality. If an article is published in a non-peer-reviewed journal, it may still be of high quality, but you can't assume that as easily.
3. Note that since the answers to the questions from Table 6.1 were exact quotations from the article, I also paraphrased to produce the abstract shown in Table 6.3.

SUGGESTIONS FOR FURTHER READING

Bauer (2003)—A guide to “the great books,” the first four chapters provide an excellent discussion of how to read scholarly texts critically.

Bean (1996)—Chapter 8, “Helping Students Read Difficult Texts,” is addressed to instructors teaching critical reading, but students can benefit from it as well.

Locke et al. (1998)—Monograph on understanding scholarly writing. The best parts are Chapter 4 on reading research reports, Chapter 5 on taking notes, and Chapter 7 on critical reading per se.

Ruggiero (1998)—Introduction to critical thinking. Chapter 4 provides thoughtful suggestions for critical reading.

Schroeder et al. (1985)—Useful survey of how to understand scholarly articles in psychology, but applicable to studies in the social sciences more generally.

Trelogan (2001)—Helpful web-based introduction to critical thinking and reading.

Wyrick (1994)—Chapter 9, on the economic content of published research, is quite good though it emphasizes the format of economic writing, rather than the argument.

EXERCISES

1. Find a scholarly article on your research topic. Read it critically, and using the article write down answers to the questions in Tables 6.1 and 6.2.
2. Write an abstract for the article you used in Exercise 1.
3. Write a critical review of the article.
4. Choose two articles on the same topic but with conflicting conclusions. Write a critical review of the two articles, and explain which article's conclusion you found the most persuasive and why.
5. Prepare an annotated bibliography for your research topic. The bibliography should include at least three sources: an article from a scholarly journal, a book, and an article from the Internet. Each item should have the three elements required for a critical review:
 - The complete bibliographic citation in the citation style required by your instructor;
 - A brief summary of the argument; please note if the source is theoretical, empirical, or a survey; and
 - An assessment of the quality of the item.

Reading and Evaluating a Theoretical Article

Reading and understanding a theoretical journal article is a little different than doing so for an empirical article. Typically, in a theoretical article, the analysis is more abstract or more in depth, which makes reading and understanding even more challenging. Before continuing, you may wish to read Chapter 7 for help in understanding theoretical analyses. Try reading the following theoretical articles, which I have annotated. In recent years, purely theoretical articles have become highly mathematical. For that reason, the following selections are taken from earlier decades.

Sample Theoretical Journal Articles

Krugman, Paul. 1979. Increasing returns, imperfect competition, and international trade, *Journal of International Economics* 9 (November): 469–479.

This article demonstrates that international trade can result from economies of scale, even when both countries have identical preferences and factor endowments. This article offers much to commend it to researchers new to theoretical studies. First, it is well written; even the analytical parts are clearly explained verbally. Second, though the details of the analysis require an understanding of constrained optimization using calculus, the analysis itself should be readily understandable to anyone who has taken intermediate theory. More precisely, the analysis consists of utility and profit maximization. Finally, the article is short, less than ten pages long.

Becker, Gary. 1962. Irrational behavior and economic theory, *Journal of Political Economy* 70 (February): 1–13.

This article established that maximizing behavior is sufficient but not necessary for obtaining downward-sloping demand curves. Indeed, several forms of irrational behavior also lead to downward-sloping demand curves as long as economic agents are subject to a budget constraint. The article uses graphs and logic but no higher math.

Tobin, James. 1956. The interest elasticity of the transactions demand for cash, *Review of Economics and Statistics* 38 (September): 241–247.

This classic article demonstrated that the transactions demand for money is a function of interest rates as well as income. Previously, only the speculative demand for money was thought to be interest sensitive.

Leibenstein, Harvey. 1950. Bandwagon, snob, and Veblen effects in the theory of consumers' demand, *Quarterly Journal of Economics* 64 (May): 183–207.

This study makes a graphical and verbal argument for nontraditional effects of price on quantity demand—for example, that high prices can induce consumers to buy because of the snob effect, which is shown by a shift in the demand curve.

Theorizing or Conceptualizing the Research

"Man, because he is endowed with reason by which he is able to perceive relationships, sees the causes of things, understands the reciprocal nature of cause and effect, makes analogies, easily surveys the whole course of his life, and makes the necessary preparations for its conduct."

CICERO

In this chapter, we explore how scholars perform the theoretical or conceptual analysis of economic research. The result is often called the theoretical model of one's research. As we observed in Chapter 2, in a real sense this is the crux of the research project: where theory is applied to shed light on the problem being studied. Many scholars have noted that developing the conceptual framework is difficult for novice researchers. For one thing, it is the most abstract aspect of the research process, requiring both analysis of the research issue or problem (in the sense of taking it apart to understand it) and synthesis of an appropriate conceptual framework to explain it. It also requires that scholars have sufficient knowledge of the appropriate economic theory on which to build. Let's see if we can explain it.

The underlying theme of this chapter is what it means to “apply” theory to a research topic. We will begin by discussing what theorizing means in general terms. We then explain the range of ways in which researchers theorize, from narrative reasoning to mathematical reasoning. We also provide examples to make this difficult topic as concrete as possible. Next, we present a shortcut that scholars commonly use: modifying an existing theoretical model rather than creating one from scratch. Finally, we identify the characteristics of a good research hypothesis, which is the product of the theorizing process.

What Does It Mean to Apply Theory to a Research Topic?

An economic research project is one in which economic theory is applied to an issue so as to derive insights about it. What exactly does it mean to “apply” economic theory to an issue? Recall the major theories you learned in your principles of economics classes. On the micro side, the major theories include supply and demand, production and cost, the theories of the firm, and consumer behavior. On the macro side, the major theories are aggregate demand and supply, consumption expenditure, investment demand, and money demand. When you apply a theory to an issue you ask yourself, “Can this question be expressed in terms of one of these theories?” In other words, you ask whether your topic is related to demand, supply, production, and so on.

Let’s start with a straightforward example. Suppose your research question involves retail diamond jewelry sales. Can you articulate this as an example of the theory of demand? The theory of demand says that the quantity demanded of some item is influenced by the price of the product; by the consumer’s income, tastes, and preferences; and by the prices of related goods. If we treat diamond jewelry sales as reflecting the demand for diamond jewelry, what would be the analogous explanatory variables? These might include diamond prices, income, wedding engagement rates as an indicator of tastes, and prices of other types of jewelry. We can conclude that the theory of demand can likely be used to analyze retail diamond jewelry sales. Note that whether the theory actually works in practice will be determined by empirical testing.

What Is Theorizing?

If after you’ve given it some thought, it is not obvious what economic theory should be applied to your research, you may need to do some formal theorizing about your topic. Theorizing is the process of brainstorm-

NOTES FOR NOVICE RESEARCHERS

Applying Economic Theory

As you review the literature for your research project, you should ask yourself for each study that you read what economic theory is being applied.

ing about an issue so as to identify the logical connections that explain the issue. The result of the process is a theory that analyzes the research question, and in particular provides an answer to the question in the form of the research hypothesis.

When you theorize, you ask three questions:

1. What are the essential concepts involved in the problem being researched?
2. How are the essential concepts related?¹
3. What implications or predictions can be drawn from these relationships?

Theorizing involves constructing a conceptual or theoretical argument. As we noted in Chapter 4, this involves deductive reasoning. As you construct this argument, you will need to think both deeply and broadly about the concepts. Let’s work through a simple example, and then we’ll look in more detail at the process.

Suppose we are investigating the extent to which the U.S. macroeconomic slowdown was caused by the decline in the stock market. How do we “conceptualize” this research question? What are the major concepts involved? The initial two concepts are the economic slowdown and the decline in the stock market. What does economic theory suggest might cause a slowdown in the economy? One answer would be a decrease in aggregate demand. What could cause such a decrease? Since aggregate demand is the sum of consumption expenditure, investment expenditure, government expenditure, and net export expenditure, a decrease in any of those components would cause a decrease in aggregate demand. Let’s consider how each of these might be related to the other major concept—a decline in the stock market. Theory tells us that though consumer spending is primarily dependent on disposable income, it also depends on household wealth. If a decline in the stock market makes individuals feel less wealthy,

it may well cause a decrease in consumer spending. A decline in the stock market could have both a direct and indirect effect on investment spending. To the extent that businesses obtain investment financing by issuing shares of stock, a decline in the market makes that more difficult. That is the direct effect. The indirect effect may be even stronger. If the stock market decline reduces consumer spending, it also reduces the demand for business output and therefore business investment.

Let's restate what we have just described.

First, we asked, given the major concepts in the problem, what economic theory or theories can shed light on them? In this case, it was theories of consumer spending and investment demand. Then, we asked, how could one apply those theories to the current research problem? Can we determine a logical connection between the stock market decline and the macroeconomic slowdown? Yes.

What are the implications of that hypothetical story? In other words, what does this theory predict? If the logic is correct, then a decline in the stock market should cause an economic slowdown as a result of reduced investment and consumer spending. This implication can be tested.

Notice the distinction between existing economic theory and the resulting theory that is developed for a specific research project. In the current example, the existing theory is the theory of consumer spending; in the earlier example, it is the theory of demand. In the current example, the resulting theory is the effects of the stock market on consumer spending; in the earlier example, it is the demand for diamond jewelry. In each case, the specific theory is an application of the existing, more general theory.² For example, any introductory economics student knows that the general theory of demand postulates a negative relation between price and quantity demanded. For the specific research project described earlier, the relevant question could be, how sensitive is the demand for diamond jewelry to changes in the price?

Narrative Reasoning

You can carry out a theoretical analysis in a variety of ways, from the less abstract to the more abstract. At one end of the spectrum is the use of a narrative reasoning process. At the other end of the spectrum is the use of formal mathematical methods. Let's examine these different approaches.

Remenyi et al. (1998, 129) argue that "when one attempts to develop models of the world, these start as narrative descriptions within which the imagination is allowed to range freely and widely over many possibilities." The way to begin this process is by creating what they call a "primary narrative," that is, a document that gives a detailed description of the research

topic, based on everything the researcher has been able to find out on the subject. The next stage is "concept creation," in which you review the primary narrative so as to identify the essential concepts. The third stage is to write a "higher order narrative," a revised version of the primary narrative that focuses on the concepts identified in the previous steps and ignores secondary details. The next stage is to examine the rewritten narrative to identify possible relationships between the concepts. The final step is to postulate hypotheses from the theoretical relationships. Though brainstorming is the strength of the narrative approach, Remenyi et al. observe that "The narratives are of course constrained by currently accepted knowledge and theories."

Though it isn't necessary to follow this exact procedure, it does give you some useful ideas about how to proceed. Other scholars have suggested some additional questions to consider as you conceptualize your research. For example, Williams (1984) proposes that scholars think about the issue or problem behind the research question, about what is known about the sources of the problem, and about what variables are relevant to the analysis of the problem. Ethridge (1995) reminds us to think about how other researchers have conceptualized problems similar to yours.

Mathematical Reasoning

Theoretical economists have traditionally used mathematical reasoning to conceptualize their research, and specifically to formulate hypotheses. Indeed, the main behavioral relationships familiar to students, including demand curves, supply curves, and production and cost functions, were all derived mathematically. Though we differentiate between mathematical and narrative theorizing, one can make a case that all research begins at a narrative level. Highly abstract theories tend to start life as "stories" prior to being formally fleshed out in terms of mathematics.

Chapter 4 defined deductive reasoning as reasoning that starts from one or more assumptions and derives a specific conclusion from them. This suggests a two-step process, as follows. First, identify the relevant economic assumptions for the problem at hand. This is analogous to the first three stages in the narrative approach discussed earlier. Second, use mathematics to manipulate the assumptions so as to derive a conclusion or hypothesis. This is comparable to the last two stages in the narrative approach. Let's illustrate this process with a series of examples.

Example 1: Optimizing Models

Economists employ two types of approaches to mathematical theorizing. The first approach is called **optimizing models**, and the second is called

ad hoc models. Optimizing models are based on the assumption that the behavior of economic agents is driven by their attempts to maximize utility or profits or to minimize costs.

Suppose our research question asks what factors determine how much output Acme Widget Company will supply? We begin by assuming that Acme will choose the level of output that will maximize profit. We know that profit is the difference between revenue and costs. Suppose that Acme's revenue and cost functions are the following:

$$\text{Total Revenue} = P * Q$$

$$\text{Total Costs} = 4 * Q^2$$

Thus, our problem has four assumptions:

1. The assumption that Acme will choose the level of output to maximize profits;
2. The definition of profit as total revenue minus total cost;
3. The definition of total revenue as $P * Q$; and
4. The definition of total cost as $4 * Q^2$

The next step is to mathematically manipulate the assumptions to deduce a conclusion.

By substituting the definitions of total revenue and total cost into the definition of profit, we obtain Acme's profit function:

$$\text{Profit} = P * Q - 4 * Q^2 \text{ or Profit} = f[Q] \text{ for short.}$$

This profit function typically has the shape of a parabola or "hilltop." If we were managing Acme, we could experiment with different levels of Q to determine the output level with the highest profit. We can more systematically solve this type of optimization problem by using calculus. Conceptually, we wish to find the "top" of the profit function, where the slope of the profit function is zero. We know that mathematically we can find the slope of a function at any point by calculating the "derivative" of the function. If we want to know where the slope equals zero, we can set the derivative equal to zero and solve for Q . In this case, the slope of the profit function, $d\text{Profit}/dQ = P - 8Q$. Setting this equal to zero and solving for the desired Q yields:³

$$0 = P - 8Q \text{ or } Q = .125 * P$$

Thus, our mathematical operations on the problem lead us to the conclusion that Acme's supply curve should be an upward-sloping function of the price of Acme's product. This conclusion is based on the positive coefficient (+.125) on P . This is a testable hypothesis, derived by manipulating the assumptions of the model.

Example 2: Ad Hoc Models

Ad hoc models are models that are not derived from optimizing principles. Rather, the hypothesized relationships come from common sense or experience. The following is an example of a hypothesis that is derived by manipulating an ad hoc model.

Suppose the Bush Administration has proposed an economic stimulus package of \$100 billion worth of spending increases. Our research question asks: What effect will the stimulus package have on GDP? Will it be sufficient to correct the slowdown? Let's assume that the U.S. economy can be modeled by a simple Keynesian model consisting of three equations:

An identity to define aggregate demand: $AD = C + I + G$

A behavioral equation to define consumption expenditure:

$$C = a + b * Q$$

An equilibrium condition: $AD = Q$

The consumption function is what makes this an ad hoc model. We did not obtain this equation as a result of trying to maximize or minimize some function. Rather, we skipped the optimizing and, based on experience, simply proposed that consumption expenditure should be positively related to income or GDP. We further assume that investment expenditure is given and that the parameters of the consumption function (a and b) are known.

Thus, we have five assumptions:

1. The assumption that the U.S. economy can be adequately modeled by three equations;
2. The definition of aggregate demand as the sum of C , I , and G ;
3. The assumption that consumer spending can be modeled by the behavioral equation given above, where the parameters are known;
4. The assumption that investment expenditure is given; and
5. The equilibrium condition.

Using algebra, we can solve the model for Q :

$$Q = [1/(1 - b)] * (a + I + G)$$

Remember that we wish to determine how the increase in government spending (dG) will affect the level of GDP (dQ). Using some simple calculus, we can see that $dQ = [1/(1 - b)] * dG$, again assuming that a and I are constant so that da and dI are both zero.

Suppose b , which represents the marginal propensity to consume, has a value of 0.75. In other words, when an individual's income increases by one dollar, he or she spends \$0.75 more. Then, from the equation just

given, $1/(1 - b)$ equals 4. Thus, a \$100 billion increase in government spending should, assuming this theory is correct, cause a \$400 billion increase in GDP. This is a testable hypothesis that was deduced from the mathematics.

A Commonly Used Shortcut: Modifying an Existing Model

Though it is always useful to go through the theorizing process just described, rarely do economists create an entirely original model from scratch. We mentioned in the beginning of the chapter that oftentimes researchers can readily apply a common economic theory to a research problem. These theories include demand and supply functions, production and cost functions, aggregate production functions, and consumption, investment, and money demand functions.

Often researchers take an existing model, which has already been applied to the topic they are interested in, and modify it in some way that seems an improvement over the original. Recall from Chapter 3 that one purpose of the literature survey is to avoid having to reinvent the wheel, and instead build on the work of scholars who came before. The literature survey often shows what has been a helpful way to approach a research problem, as well as what hasn't worked. So, starting with an existing model that was at least partially successful and then modifying that model to improve it is a time-honored approach.

This can be done either by going through the formal optimizing process or by adopting a more ad hoc approach. The ad hoc way is appropriate if you know what the outcome is going to be. For example, if you apply a very well-known theory, say, the theory of demand, it makes little sense to work through the mathematics of utility maximization so as to derive a demand curve, since all economists know that the result will be that quantity demanded should be negatively related to the price. Let's illustrate this with two examples.

Example 3: Intertemporal Utility Maximization with Money

Economists going back at least to Irving Fisher have known that economic agents' decisions about consumption versus saving can be analyzed as an application of intertemporal utility maximization. Indeed, this was the idea behind Modigliani's Life-Cycle Hypothesis for Consumption Expenditure. In that model, income in each period is allocated between two items: consumption, which provides immediate satisfaction, and saving, which earns interest and future satisfaction.

The Existing Theory: Consumption Today versus Consumption Tomorrow We can first look at utility as the choice between consumption today versus consumption tomorrow.

$$U = f(C_0, C_1)$$

We can then maximize this, using the Lagrangian technique, subject to the following budget constraint:

$$\lambda(C_0 - Y_0 - (Y_1 - C_1)(1 + r))$$

The following set of equations shows the maximization subject to consumption today, consumption tomorrow, and λ .

$$\partial U / \partial C_0 = \lambda$$

$$\partial U / \partial C_1 = \lambda / (1 + r)$$

$$\partial U / \partial \lambda = C_0 - Y_0 - (Y_1 - C_1)(1 + r)$$

We can look at the marginal rate of substitution here by the following equation:

$$(\partial U / \partial C_0) / (\partial U / \partial C_1) = \lambda / (\lambda / (1 + r)) = 1 + r$$

Therefore, the choice between consumption today versus consumption tomorrow depends on the interest rate.

For a specific functional form, we use a log linear function, as follows:

$$\ln C_0 + \ln C_1 - \lambda(C_0 - Y_0 - (Y_1 - C_1)(1 + r))$$

Maximizing this function, we get the following set of equations:

$$\partial U / \partial C_0 = 1/C_0 - \lambda$$

$$\partial U / \partial C_1 = 1/C_1 - \lambda(1 + r)$$

$$\partial U / \partial \lambda = C_0 - Y_0 - (Y_1 - C_1)/(1 + r)$$

Solving the first-order conditions for C_0 implies that consumption is proportional to wealth.⁴

$$C_0 = aW_0$$

Modification of the Existing Theory: The Addition of Money Suppose we wanted to add another item to the basic model, namely, money. We would sketch out the basic model in the traditional way, but then add real money to the utility function.⁵

$$U = f(C_0, C_1, M_0, M_1)$$

In accordance with our income equation the budget constraint will be:

$$\lambda(C_0 - Y_0 - \Delta M_0 - (Y_1 - C_1 - \Delta M_1)/(1 + r))$$

Maximizing this function yields the following set of equations:

$$\partial U / \partial C_0 = \lambda$$

$$\partial U / \partial C_1 = \lambda / (1 + r)$$

$$\partial U / \partial \Delta M_0 = -\lambda$$

$$\partial U / \partial \Delta M_1 = \lambda / (1 + r)$$

$$\partial U / \partial \lambda = (C_0 - Y_0 - \Delta M_0 - (Y_1 - C_1 - \Delta M_1) / (1 + r))$$

In addition to the marginal rates of substitution found in Example 3, we also find the following:

$$(\partial U / \partial C_0) / (\partial U / \partial \Delta M_0) = -1$$

$$(\partial U / \partial C_1) / (\partial U / \partial \Delta M_1) = 1$$

$$(\partial U / \partial \Delta M_0) / (\partial U / \partial \Delta M_1) = -1 / (1 + r)$$

We can see from this that the marginal rate of substitution between money today versus money tomorrow is based on the interest rate.

As a result of this utility maximization function, we can see the following:

$$C_0 = aW_0$$

$$M_d = bW_0$$

In other words, money demand in any period should be proportional to total wealth, just as consumption is.

Example 4: Ad Hoc Example

The economics literature on the effects of parental work hours on children's cognitive development typically applies the notion of an educational production function. The idea is that just as a firm's output is produced by combining capital, labor, and other inputs, one can treat cognitive development as being "produced" by combining inputs of parental time and quality, family income, and so on. To date, the literature has not shown a consensus about whether maternal employment adversely affects children's cognitive development. Suppose you hypothesize that the reason for the mixed results is the failure to control for the quality of child-care parents use. It would be straightforward to add child-care quality as an additional input to the educational production function, treating it analytically as any other input.

What Makes a Good Research Hypothesis?

The product of the theorizing process should be a research hypothesis. In Chapter 4, we noted that an argument could be considered both the principal assertion and also the entire package of assertion and supporting evi-

NOTES FOR NOVICE RESEARCHERS

Designing the Research Project

Novice researchers often have difficulty determining how to turn a research problem into an original research design. An original research project does not require a truly original approach. Rather, it is common practice to modify an existing approach. This can be done in several useful ways. For example, *we can take the same theoretical model as was employed in a previous study*, but use:

A different method of empirical testing, or

A different data set (e.g., a different country, or a different time period).

Alternatively, *we could take a model that has been used to study some other topic* and apply it to the topic we are interested in. Then again, *we can modify the theoretical model* in some way, as we have described in this chapter. For example, we could take Economist A's basic model, but use Economist B's method for doing some aspect of A's model.

dence. In the same way, one can think of the hypothesis and the theoretical explanation as the answer to the research question. The hypothesis is the crux of the conceptual or theoretical analysis. To assess a theoretical argument, we test the hypothesis. One of the most important factors leading to a successful research project is a good research hypothesis. Ethridge (1995) suggests that a good research hypothesis must have the following characteristics:

1. It should be stated clearly and specifically in a way that can't be misinterpreted.
2. It must be able to discriminate clearly from alternative hypotheses.
3. It must be capable of being proved false.
4. It should be empirically testable (and nontrivial); that is, there must be reasonable statistical means and reasonable data available for testing it.
5. It must be derived from the theoretical analysis; otherwise, any statistical results will lack validity, demonstrating at best correlation rather than the causation that we seek. Note that either an optimizing or an ad hoc approach can be used here, as long as the hypothesis is linked to the underlying theory.

Let's illustrate with the diamond jewelry example. The hypothesis can be stated clearly that the demand for jewelry depends negatively on the price (Characteristic 1). The alternative hypothesis would be that the demand does not depend negatively on the price. This is a clear distinction (Characteristic 2). Either the hypothesis is true or false (Characteristic 3). The hypothesis can be tested using regression analysis and data obtained from the industry and public sources (Characteristic 4). Finally, as we showed earlier, the hypothesis is derived from the theory of demand (Characteristic 5).

SUMMARY

- When you apply economic theory to a research question, you ask, "Can this question be expressed or interpreted in terms of a basic economic theory such as supply, demand, production, cost, and so on?"
- Theorizing is the process of brainstorming about an issue to identify the essential concepts involved and the logical connections between those concepts.
- The result of theorizing should be a hypothesis that provides a tentative answer to the research question.
- Economic researchers theorize using some combination of narrative and mathematical reasoning.
- Theorizing is an exercise in deductive reasoning—for example, where economic assumptions are manipulated mathematically to derive a prediction or conclusion.
- It is common practice for researchers to modify an existing model in the literature rather than creating an entirely original model from scratch.

NOTES

1. Wyrick (1994) notes that the validity of these proposed relationships is strengthened if they draw on the findings of other researchers.
2. In applying the general theory, one might also ask if there are any institutions or other real-world constraints that are assumed away in the theory but whose presence in the real world has a bearing on our analysis? If so, what are the effects likely to be (e.g., rent controls)?
3. This result was obtained using the power rule: for $Y = aX^n$, $dY/dX = n \cdot aX^{n-1}$.

4. See Miller et al. (1974), p. 95.
5. This modification is from an undergraduate research paper, Fallen (2000).

SUGGESTIONS FOR FURTHER READING

- Epstein and Kernberger (2005), Chapters 4 and 5*—Excellent introduction of the logic behind economic theories and models. Thoughtful examples fit right in with the discussion in this chapter.
- Ethridge (1995), Chapter 8*—Provides a nontechnical explanation of the theorizing process.
- Remenyi et al. (1998)*—Offers a detailed explanation of the narrative reasoning process.
- Wyrick (1994), Chapter 10*—Provides another perspective on theorizing, addressed to undergraduate researchers.

EXERCISES

1. What economic theory can be applied to the following topics?
 - The effects of foreign direct investment on economic growth in developing countries.
 - Are automobile sales more sensitive to changes in price or financing?
 - The effects of free agency on salaries in baseball.
 - Housing prices in the [fill in a city] area.
2. Choose three studies from your literature survey that you found to be relevant to your research topic. Identify the economic theory that each study applies to its research question.
3. When you decide to apply an economic theory like supply or demand to your research as we described in the beginning of the chapter, are you choosing an optimizing model or an ad hoc model?
4. *Example of using an optimizing model to theorize:* Determine the demand for advertising by adding advertising as an input in the production function and then minimize cost, subject to an output constraint.
5. *Example of using an ad hoc model to theorize:* The macro economy has been modeled frequently using the Cobb-Douglas production function: $Q = A \cdot L^a \cdot K^b$, where Q = GDP, L = labor input, K = capital input, and A is technical change. This production function has great appeal, since the

coefficients on labor and capital (a and b) are interpreted as the marginal products of labor and capital. Suppose one wishes to analyze the effects on A of explicitly treating energy as a factor of production. Explain how to do that.

6. Determine an appropriate theory to apply to your research question. Use either a common economic theory or modify an existing theory from the literature on your topic. Explain why the theory is appropriate for your research question.

Locating (and Collecting) Economic Data

"It is a capital mistake to theorize before one has data. Insensibly one begins to twist facts to suit theories, instead of theories to suit facts."

SIR ARTHUR CONAN DOYLE

This is the first of two chapters on locating and collecting economic data. In this chapter, we focus on how data are constructed and where they may be found, both in primary and secondary sources. In the next chapter, we explain how to compile a data set for your own research project.

A key part of any empirical research project is collecting and manipulating data. It is never too early to begin looking for potential data sources. After all, a research project can be hamstrung by inadequate data just as easily as by a lack of a clearly conceived, testable hypothesis. In fact, it is not uncommon for novice researchers to invest a great deal of time and effort developing a research project only to discover that data aren't available to adequately test the hypothesis. Don't let this happen to you.

Even though this book describes the research process sequentially, in practice, many steps can be performed in parallel or at least in overlapping steps. Once you have settled on a research question, it is wise to begin looking for usable data. To begin, you should look for data in the general area of your research (e.g., macro data, international trade, financial data, etc.). As you review the literature on your topic, pay attention to the data sources previous studies have used. Think about their strengths and weaknesses. Would they work for your research?

Once you have sketched out your conceptual or theoretical model, you should delve into your data search in earnest. What variables do you need data for in order to test your research hypothesis? Oftentimes, researchers must simplify their theoretical model because the data are inadequate. Either there are insufficient data points for one or more variables, or a specific variable they need is not available. Before we get into the details of developing your own data set, to be discussed in Chapter 9, we first need to examine the process of data creation and the structure of published data.

Data Creation

In Chapter 1, we noted that many people think that knowledge consists of facts that can be plucked from a tree in the forest, once that tree is discovered. By contrast, we have argued that knowledge is created; that knowledge is the interpretation of facts.

In the same way, people tend to think of statistics as facts. This is misleading or wrong. The vast majority of data are constructed rather than collected. Government or private organizations that do a serious job of constructing statistics make decisions throughout the process that affect the quality of that data for any given purpose.

Steps in Data Construction

Best (2001) identifies three steps in the construction of a data series:

1. Defining the concept;
2. Deciding how the concept will be measured; and
3. Determining how to define the sample on which the data will be based.

Let's examine these steps to get a better sense of what published data really mean.

Suppose we are interesting in finding income so we can test the theory of consumer behavior. What is the basic consuming unit? There are two commonly used possibilities: households and families. If we choose families, our concept of income will be average family income. If we choose households, it will be average household income. If you compare the two variables, they will be similar but not the same.

This illustrates an important point about data. Every data series is constructed for a specific purpose. You should realize that that purpose may not be the same as yours. As a consequence, a given data series may not be defined or measured in a way that best matches your needs. Unless you are going to construct your own data, you will have to live with this problem.

However, you can minimize this problem by doing your best to discover exactly where the data come from and what they actually measure.

Suppose we've chosen to define our variable as average family income. How is that to be measured? What will be included? Should we include labor income only, for example, wages and salaries, or capital income, for instance, interest, dividends, capital gains, too? Should we include in-kind payments, such as food grown on farms and consumed by the farmers? It's not sufficient to decide you want "total" income. You still need to define what is included.

How are we going to measure average family income? What measure of the average should we use? Unless the data are approximately normally distributed, which income data tend not to be, you can get substantially different values for the mean, median, and mode. Thus, for income data one tends to use the median rather than the mean as a measure of central tendency.

Sample Data

Most social science statistics are based on sample data rather than populations. Therefore, the data that are published are extrapolated from samples. For example, average family income is not, in fact, the average income of all families; rather, it is the average income of the families in the sample.

Only if the sample is random, and thus truly representative of the population, will the sample statistics correctly measure the population. A truly random sample is difficult and expensive to acquire, so typically samples are less than truly random. As a consequence, bias creeps into the results. As such, you should think of all data as *estimates* rather than facts.

Best (2001, 161) sums it up well: "Every statistic must be created, and the process of creation always involves choices that affect the resulting number. . . . No statistic is perfect, but some are less imperfect than others."

The Structure of Economic Data

When we discussed surveying the literature on a research topic, we mentioned that knowing how the economics profession organizes its literature can help scholars find information quickly and efficiently. The same thing can be said for data.

It is important to differentiate between *those organizations that collect or produce data* and *those that publish it*. This is the same distinction as primary versus secondary sources of information. Let's talk first about the sources of data, and then we can figure out the best places for users to obtain that data.

Statistics are not collected at random. Rather, they are typically the result of a specific data collection effort or process. The product of this process is a specific data set that includes a collection of certain variables. The characteristics of data sets are described in the next section. The section that follows provides an overview of the various agencies that collect the data and the major data sets they produce. Then, we will explain the most important of those data sets in detail.

Characteristics of Data Sets

Survey data comes in three forms: time series, cross-section, and longitudinal (or panel) data. **Time-series data** gives different observations or data points on the same variable at different points across time. For example, a time series could consist of annual data for U.S. GDP over the time period 1950-2000. **Cross-section data**, by contrast, gives different observations of a comparable variable at the same point in time. For example, a researcher could obtain a cross section of data that consisted of average disposable personal income across the fifty states in the United States. Finally, **longitudinal data** take a cross-section sample and follow it over time.¹ For example, a sample of family income for the same ten families over five years would be a longitudinal data set. Note that all three of these examples have a sample size of fifty observations, but the basis for these observations is different. Longitudinal data are an example of a **micro data set**, meaning that the data points or observations are of individual economic agents such as individuals, households, or firms. Contrast this with macro data compiled at the national level, which might measure national income or total expenditure by consumers.²

Time-series data are available in different frequencies. Frequency means how often the concept is measured. The U.S. Census is conducted every ten years. U.S. GDP is measured quarterly. Unemployment is available monthly. Interest rates are available daily, and stock prices may be available hourly. Each of these variables is also available for longer time periods. For example, U.S. GDP, unemployment, interest rates, and stock prices are available at annual rates, too. Figure 8.1 lists a selection of commonly used time-series variables and the frequencies with which they are measured.

Cross-section data have a comparable concept to data frequency that involves the "unit of analysis." With time series data, the observations are taken at different points in time. By contrast, with cross-section data, the observations are measured across different nations, states, or other units of analysis. Suppose you need data for income. Average disposable personal income, for example, can be measured in each of the fifty states. Al-

Figure 8.1 Available Frequency of U.S. Time Series Data

Data Available Annually:	Fixed Reproducible Capital—Fixed-Asset Tables
Data Available Quarterly:	Gross Domestic Product and Major Components— Advance Estimate, available with a 1-month lag; Preliminary Estimate, available with a 2-month lag; Final Estimate, available with a 3-month lag. Corporate Profits Productivity Growth Employment Cost Index Flow-of-Funds Accounts U.S. International Transactions (complete balance of payments: current + capital account), available with a 2.5-month lag.
Data Available Monthly:	Personal Income and Outlays, Disposable Personal Income, Personal Consumption Expenditures, available with a 2-month lag Retail Sales Employment and Unemployment Rates Consumer Price Indices Producer Price Indices U.S. International Trade in Goods and Services (i.e., current account), 1.5-month lag. Index of Industrial Production, Capacity Utilization Rate, 1.5-month lag. Consumer Credit, 5-week lag.
Data Available Weekly:	Monetary Base, Monetary Aggregates
Data Available Daily:	Exchange Rates, Selected Interest Rate

ternatively, you might be able to obtain family income data across all counties in a particular state, or data on household income for sample households in a particular county. The researcher's decision about what frequency or unit of analysis to select can be guided by economic theory, but is often determined by data availability.

Organizations That Collect and Publish Data

A number of U.S. governmental, international, and private organizations gather economic and social statistics.³ These include the Census Bureau, the Bureau of Economic Analysis (BEA), the Bureau of Labor Statistics (BLS), the Federal Reserve, the International Monetary Fund, the World Bank, and the United Nations, as well as private organizations such as the Conference Board. Let's survey some of these.

Census Bureau

The **Census Bureau** (www.census.gov) constructs general and population statistics, as well as foreign trade data on goods and services imported into or exported from the United States. An excellent place to start looking for general statistics is the *Statistical Abstract of the United States*. Though the *Statistical Abstract* may not have the specific data you need, the references (source information) for each of its tables can give you clues about where that specific data may be found. The Census Bureau publishes several other useful references, such as the *State & Metropolitan Area Data Book*, *USA Counties*, and the *County & City Data Book*. The Census Bureau conducts the decennial *U.S. Census of Population*, which includes detailed socioeconomic data. It also conducts several lesser-known surveys, such as the *Economic Census* (every five years), the *Annual Survey of Manufactures*, *Current Industrial Reports* (annual/quarterly/monthly), and the *American Housing Survey*. Finally, the Census Bureau conducts a number of surveys on behalf of other agencies (e.g., BLS). These will be discussed later.

Bureau of Economic Analysis

The **Bureau of Economic Analysis** (www.bea.doc.gov) constructs the major U.S. macroeconomic indicators. The BEA also produces estimates of private and public capital stocks (under the heading "Fixed Assets"), GDP by industry, U.S. balance of payments, U.S. international investment position, and regional and state data. The BEA has several major data collections, most prominently the *National Income & Product Accounts* (www.bea.doc.gov/bea/dn/nipaweb). It includes detailed estimates of GDP and its components (e.g., personal consumption expenditure, investment expenditure, government expenditure, and net exports), as well as price indices for each. The report also provides a detailed breakdown of national income. The BEA publishes the monthly *Survey of Current Business* (<http://www.bea.doc.gov/bea/pubs.htm#SCB%20Table>), which provides notice of revisions of each of its products as well as detailed discussion of data sources and methodology.

Bureau of Labor Statistics

The **Bureau of Labor Statistics** (www.bls.gov) constructs data related to employment issues, productivity, and prices (e.g., the Consumer Price Index, the Producer Price Index, the Employment Cost Index, and components of each). The BLS produces a number of major data collections, which will be discussed later in this chapter, including the *Current Population Survey*, *Current Employment Statistics*, the *Consumer Expenditure Survey*, as well as the *National Longitudinal Surveys*. The BLS also publishes the *Monthly Labor Review*, which plays a role for the BLS similar to that of the *Survey of Current Business* for the BEA.

The Federal Reserve

The **Federal Reserve** (www.federalreserve.gov) constructs principally financial data: interest and exchange rates, money stocks and components (e.g., bank reserves), debt measures, bank assets and liabilities, official reserves, household assets and liabilities, and corporate debt. These data are available at <http://www.federalreserve.gov/rnd.htm>. The Fed also conducts the *U.S. Flow of Funds Accounts* and the *Survey of Consumer Finances*. Though much of the data from the Federal Reserve System is produced by the Board of Governors (BOG), researchers should not neglect the data available from the regional Federal Reserve banks, especially if one is interested in regional economic research. Some of these data are not available from the BOG. For example, the Atlanta Federal Reserve bank computes a trade-weighted foreign exchange value for the U.S. dollar, based on the fifteen-largest trading partners of the United States. Also, the St. Louis Fed has an easy-to-access collection of U.S. macro data (called FRED II).

International Agencies

A number of international agencies also collect economic and financial data of use to researchers. These agencies include the International Monetary Fund, the World Bank, the Organization for Economic Cooperation and Development, EuroStat, the Asian Development Bank, and the Inter-American Development Bank. A number of United Nations' agencies also publish data. Most of this data, however, is available in electronic formats only via expensive subscriptions. However, it is fairly widely available in print formats in many college and university libraries.

Major Primary Data Collections

In this section, you can find more detailed explanations of the major data collections mentioned previously, as well as other major primary-data collections. A more comprehensive survey of the major data collections may be found in Maier (1999).

U.S. National Income and Product Accounts

The National Income and Product Accounts (NIPA) are the official national accounts (i.e., macroeconomic statistics) of the United States and can be found at (<http://www.bea.doc.gov/bea/dn/nipaweb/index.asp>). This is the source of all U.S. macroeconomic data. The accounts include real and nominal data on national income and product; personal income and expenditures; government receipts and expenditures at the federal, state, and local levels; foreign transactions; saving and investment; income and employment by industry; and price indices for nearly all these categories. This data set offers an amazing amount of detail. For example, Table 2.6, "Personal Consumption Expenditure by Type of Product," indicates that U.S. residents spent \$80.6 billion on higher education in 2000, but only \$62.1 billion on dental services. Annual data are available back to 1930, and quarterly data (for the major series) are available back to 1952. Much of the data are available online and are easily downloadable in spreadsheet format. For additional information, see Maier (1999), Chapter 7.

U.S. Flow of Funds Account

The Flow of Funds Account (<http://www.federalreserve.gov/releases/Z1/>) collects data on financial flows across the U.S. economy. It includes detailed information on levels of and changes in assets and liabilities in each sector of the economy. Annual data are available back to 1945, quarterly data are available back to 1952, and monthly data for the major components of domestic nonfinancial debt are available back to 1955.

U.S. Balance of Payments Accounts and International Investment Position of the U.S.

The international accounts of the U.S. (<http://www.bea.doc.gov/bea/di1.htm>) are compiled in two sets of data that correspond to a flow account and a stock (or asset) account. The flow account measures exports and imports of goods and services (the Current Account) as well as capital in- and outflows (the Capital Account). These data are available quarterly for the aggregate data that measure U.S. balance of payments vis-à-vis the rest of the world. Quarterly data are also available for U.S. international transactions with Western Europe, the European Union, the United Kingdom, Eastern Europe, Canada, Latin America, Japan, and Australia. Annual data are available for trade with Belgium-Luxembourg, the Netherlands, Germany, France, Italy, Mexico, Venezuela, and South Africa. The aggregate data and some of the country data are available back through 1960. The stock account data, which are collected annually, summarize U.S. holdings of assets from and liabilities to the rest of the world. More detail is avail-

able (by country and industry) for data on U.S. direct investment abroad and foreign direct investment in the United States.

U.S. Census of Population and Integrated Public Use Microdata Series

The U.S. Census (<http://www.census.gov/main/www/cen2000.html>) is the decennial measurement of the U.S. population. In addition to counting the total number of Americans, the Census collects a wide range of socioeconomic information that is disaggregated by location, from the national level all the way down to zip codes. The Minnesota Population Center's Integrated Public Use Microdata Series (IPUMS, <http://www.ipums.umn.edu/>) is comprised of twenty-five "high-precision" samples of the U.S. population from U.S. censuses since 1850. The data include economic and social variables. For additional information, see Maier (1999), Chapter 2.

Current Population Survey

The Current Population Survey (CPS, <http://www.bls.gov/cps/home.htm>) is a monthly survey of a sample of U.S. households designed to collect data on employment, unemployment, and individuals not in the labor force. Its data are classified by a number of demographic characteristics, including age, gender, race, occupation, industry, and hours of full- or part-time employment. Some data are available on the work experience and educational background of workers. For additional information, see Maier (1999), Chapter 9.

Current Employment Statistics

Current Employment Statistics (CES, <http://www.bls.gov/ces/home.htm>) is a monthly survey of payroll records of a large sample (more than 500,000) of U.S. businesses. It collects data similar to that of the CPS, but the source is employers rather than households. Specifically, the CES collects detailed data on employment (i.e., number of employees), weekly work hours, and hourly and weekly earnings of nonfarm workers by industry. The data are organized in two ways: Ownership (private versus public employer, and if the latter, then federal, state, or local government) and Industry (using the North American Industry Classification System (NAICS)). Data are available at the national level, as well as by state and metropolitan area.⁴ National data for employment are available back to 1919. National data for work hours are available back to 1947; national data for hourly earnings are available back to 1962. Data for states and metropolitan areas are available back to the 1980s.

The Economic Census

The Economic Census (<http://www.census.gov/epcd/www/econ97.html>) provides a detailed snapshot of the production side of the U.S. economy. It covers all the production sectors of the economy, except for agriculture and government (which are addressed in separate surveys). These include: mining; utilities; construction; manufacturing; wholesale and retail trade; transportation and warehousing; information; finance; insurance; real estate; rental and leasing; professional, scientific, and technical services; management; educational services; health care; entertainment; food services; and others. Beginning with the 1997 Economic Census, data are organized under the NAICS industry classification system rather than the previous Standard Industrial Classification (SIC) system. The Economic Census is a survey of "establishments," which are defined as individual business units at a single location. Thus, a single company (e.g., Pizza Hut) may include multiple establishments. Key data collected include the number of establishments, number of employees, payroll, measures of sales or output, costs, and assets. Data are available (in decreasing detail) at the national, state, metropolitan area, county and zip code levels, and also by specific industry.

Annual Survey of Manufactures

The Annual Survey of Manufactures (ASM) is a survey of a sample of manufacturing establishments organized according to the NAICS industry classification system. The survey can be found at <http://www.census.gov/econ/overview/ma0300.html>. The survey includes data on value of shipments, value added, employment, labor costs, materials costs, capital expenditures, energy consumption, and inventories.

Current Industrial Reports

The Current Industrial Reports (CIR, <http://www.census.gov/cir/www/index.html>) provide annual, quarterly, and monthly data on production and shipments of manufactured goods in selected industries, including aerospace equipment, chemicals, computer and electronic components, consumer goods, industrial equipment, primary metals, textiles, and apparel. The most disaggregated data are available at the annual level; the least at the monthly level.

American Housing Survey

The American Housing Survey (AHS, <http://www.census.gov/hhes/www/ahs.html>) is a panel survey conducted biannually at the national level and every four years for major metropolitan areas. The national panel includes approximately fifty-five thousand housing units, including single-family homes, apartments, and mobile homes. The survey also in-

cludes data on housing and neighborhood characteristics, housing costs, and income of the residents.

Consumer Expenditure Survey

The Consumer Expenditure Survey (CES, <http://www.bls.gov/cex/home.htm>) is a survey of the spending behavior of a sample of American households. The data are derived from a quarterly interview and a weekly diary that respondents fill out. The result is annual data on income and expenditures disaggregated by age, household size, and other demographic characteristics. Expenditures are broken down by major classes, including food, housing, apparel and services, transportation, health care, entertainment, personal care products and services, reading, education, tobacco products and supplies, and personal insurance and pensions. The data are available back to 1980 for some series, and 1984 for the rest. For a similar data set, see Tables 2.6 and 2.7 of the NIPA.

National Longitudinal Surveys

The National Longitudinal Surveys (NLS, <http://www.bls.gov/nls/home.htm>) are a collection of surveys, conducted by the Bureau of Labor Statistics over the past thirty years, to develop data on the labor market activities, broadly defined, of several representative groups of Americans. These are probably the longest running and certainly the best-known panel data sets in the United States. The original surveys consisted of four cohorts: Young Men, Young Women, Mature Women, and Older Men. All cohorts were constructed to be nationally representative samples of their respective groups. The surveys include data on education, training, marriage, fertility, health, income, and assets.

The Young Women's survey includes women who were ages fourteen through twenty-four when first interviewed in 1968. The Mature Women's survey includes women who were ages thirty through forty-four when first interviewed in 1967. These surveys are now conducted simultaneously in odd-numbered years. The Young Men's survey, which was discontinued in 1981, includes men who were ages fourteen through twenty-four when first interviewed in 1966. The Older Men's survey, which was discontinued in 1990, includes men who were ages forty-five through fifty-nine when first interviewed in 1966.

There have been three subsequent cohorts: "Youth 1979," "Children of Youth 1979 - Supplement to NLSY79," and "Youth 1997." The NLSY79 is a survey of men and women born in the years 1957-64; respondents were ages fourteen through twenty-two when first interviewed in 1979. The NLSY79 "Children and Young Adults" is a survey of the children of the women in the NLSY79. The NLSY97 is a survey of young men and women

born in the years 1980-84; respondents were ages twelve through seventeen when first interviewed in 1997.

The NLS79 includes data on wages, work hours, industry, and occupation. It also includes some information on spouse employment; highest level of education completed (by grade); high school GPA and DOD aptitude test (ASVAB) scores; job training; family relationships (marriage, number of children, etc.). The "Children of Youth 1979" includes data on cognitive, socio-emotional, and physiological assessments; household income; net worth; and health status. For a similar data set, see the Panel Study of Income Dynamics (PSID).

Panel Study of Income Dynamics

The Panel Study of Income Dynamics (PSID, <http://www.isr.umich.edu/src/psid/>) is a longitudinal study of a representative sample of U.S. families. The study includes data on the families as well as the individual men, women, and children in them. Begun in 1968, the data were collected annually through 1997, and biennially starting in 1999. The sample size began as forty-eight hundred families. By 2001, it had grown to over seven thousand families. Though similar to the NLS in terms of content, the PSID focuses more explicitly on families.

The PSID includes data on income sources and amounts; poverty status; public assistance in the form of food or housing; other financial matters (e.g., taxes, inter-household transfers); family structure and demographic measures (e.g., marital events, birth and adoptions, children forming households); labor market work (e.g., employment status, work/unemployment/vacation/sick time, occupation, industry, work experience); housework time; housing (e.g., own/rent, house value/rent payment, size); geographic mobility (e.g., when and why moved; where family head grew up; all states head has lived in); socioeconomic background (e.g., education, ethnicity, religion, military service; parents' education, occupation, poverty status); and health (e.g., general health status, disability).

Beginning in 1997, the PSID includes a Child Development Supplement. This survey uses a subsample of the PSID consisting of 2,394 households and 3,563 children up to twelve years old. Its purpose is to collect data, every five years, on a variety of issues impacting the cognitive and emotional development of children.

Surveys of Consumers

The Survey Research Center (<http://www.sca.isr.umich.edu>) at the University of Michigan publishes a number of indices that measure household expectations. The most prominent of these is the *Index of Consumer Senti-*

ment. These data are accessible through the "guest" login at the Survey Research Center. They are also available from the FRED II database at the St. Louis Fed.

The Conference Board (<http://www.conference-board.org>) publishes the monthly *Consumer Confidence Index*. The most recent monthly data are available free from their webpage, but past data require a subscription.

Survey of Consumer Finances

The Survey of Consumer Finances (SCF, <http://www.federalreserve.gov/pubs/oss/oss2/scfindex.html>) is a survey conducted by the Federal Reserve Board of Governors of selected demographic characteristics of U.S. families, including their income, balance sheets, and use of financial services. The survey has been triennial since 1983. For similar data on the U.S. economy as a whole, see the U.S. Flow of Funds Accounts.

Major Secondary Data Collections

In the previous section we identified the major producers of U.S. economic and social data. You can certainly obtain data from those sources, especially if you are willing to dig for it. However, it is often easier to go to a secondary source of data. Secondary sources are usually more user friendly than primary sources, which tend to be designed for expert users. Additionally, just as it may be helpful to begin a literature search with secondary sources, the same is true for data searches, especially if you are uncertain about what data are available on your topic or where those data may be found. What follows is not meant to be an exhaustive list but rather a sample of what is available. It also includes a sample of the major secondary sources of international economic and social data.

The Economic Report of the President

The Economic Report of the President (ERP, <http://w3.access.gpo.gov/eop/>) provides easy access to nearly all the macroeconomic data undergraduates need for their courses (though not for their research papers). This is a good first place to look for macro data.

Economagic

Economagic (www.economagic.com) is an excellent source of U.S. macro and regional as well as some foreign data. The advantages of Economagic are, first, that it is a very large site, covering data for more than one hundred thousand variables, and second, that it provides a single, easy-to-use interface for accessing all that data, which is produced by a variety of government agencies. The data are available in multiple formats from spreadsheets to .PDF files. The site also allows users to do some data manipulation

before downloading. For example, if the data available focus on GDP but you need growth rates, Economagic will compute them for you. More advanced features are available by subscription.

FRED II (Federal Reserve Economic Data)

FRED II (<http://research.stlouisfed.org/fred2/>) is an excellent source for U.S. macroeconomic and financial data. It also provides regional data for the region of the country serviced by the St. Louis Federal Reserve Bank. The data are easy to download.

STAT-USA/State of the Nation

STAT-USA (<http://www.stat-usa.gov/>) is a U.S. federal government-sponsored data site. It provides a single interface for obtaining U.S. macroeconomic, financial, and industry data. The primary data are obtained from the U.S. Bureau of Economic Analysis, the U.S. Bureau of Labor Statistics, and the Federal Reserve. The data are available by subscription or for free through federal depository libraries, which includes many universities. STAT-USA is not as user friendly as Economagic or FRED II. The main advantage of this site is that it includes some specialized data that are unavailable from those other sources.

Inter-university Consortium for Political and Social Research

Inter-university Consortium for Political and Social Research (ICPSR, <http://www.icpsr.umich.edu/>) is a major archive of data for social science research. It covers data on a broader range of topics, for example, survey and panel data, than the other (i.e., economics) sources discussed here. It is a good source of data for microeconomic research. ICPSR is a subscription service, but it is available to researchers at many universities.

International Financial Statistics

International Financial Statistics (IFS) is the principal data set of the International Monetary Fund (IMF). It is an excellent source of macroeconomic, international, and domestic financial data for IMF member countries. Data include exchange rates, international liquidity, international banking, money supplies, interest rates, price levels, international trade, government accounts, and GDP and its major components. Data are available in monthly, quarterly, and annual frequencies. Not all data series are available for all countries or for all frequencies. The main negative feature (from an undergraduate's point of view) is access. IFS is widely available in hard copy, but it is not available online. It is possible to purchase the data on a single-user CD-ROM, but the price (\$350 for the academic rate) makes it prohibitive for many colleges. For annual data, the *IFS Yearbook*, even in hard copy, can provide a quick fix. One additional caveat is

that since the data are supplied from member-country governments, you should not assume that they are directly comparable. For such comparisons, a better data source might be the Penn World Tables (discussed later in this section).

World Economic Outlook Database

The World Economic Outlook is a biannual (May and October) survey of world economic conditions published by the IMF. The database (http://www.imf.org/search97cgi/s97is_eng.dll/search97cgi/inetsrcheng.ini?action=FilterSearch&filter=spquery.hts&QueryText=weodb) is designed to complement the survey. It provides annual data (back to 1970) on roughly fifteen macro series, for IMF member countries. These include GDP (various measures), inflation, government budget balance (actual and structural), output gaps, net capital flows, and external debt and debt service. There are also half a dozen "world" macro aggregates, including trade volume.

Penn World Tables

The Penn World Tables (<http://datacentre.chass.utoronto.ca/pwt/index.html>) provide annual data for more than 150 countries on twenty-nine macro and related aggregates. The strength of this data set is that it offers the ability to make meaningful international comparisons between each country's data. The range of data available is 1950 through 1992. Data are not available, however, for all countries over the full range.

Joint BIS-IMF-OECD-World Bank Statistics on External Debt

The Joint BIS-IMF-OECD-World Bank Statistics on External Debt database (<http://www.oecd.org/statistics/jointdebt>) provides definitive data on nations' external debt. It includes data on debt balances and changes in debt.

Eurostat

Eurostat (http://epp.eurostat.cec.eu.int/portal/page?_pageid=1090,1137397&_dad=portal&_schema=PORTAL) is the official statistical agency for the European Union. As such, it provides a large amount of high-quality data on a wide array of subjects, including national accounts, foreign trade, and finance. Current data are available for free, but historical time series are fee-based.

OECD Main Economic Indicators and National Accounts

The Organization for Economic Cooperation and Development (OECD, http://www.oecd.org/statsportal/0,2639,en_2825_293564_1_1_1_1_1,00.html) compiles a number of excellent databases, including Main Economic Indicators and National Accounts of OECD Countries. These

cover the major macroeconomic and financial data for the thirty member countries. The major negative feature of these data sets is accessibility. Though they are widely available in hard copy, the online versions are fee-based.

A final note: If you don't find what you need in a secondary source, do consider the primary sources!

SUMMARY

- Data are constructed rather than collected.
- Most social science statistics are based on sample data.
- Survey data come in three forms: time series, cross-section, and longitudinal.
- The majority of U.S. data are produced by the Census Bureau, the Bureau of Economic Analysis, the Bureau of Labor Statistics, and the Federal Reserve System.
- Secondary data sources are often easier to use than primary data sources.
- An overview of primary and secondary data sources is given in Appendix 8A.

NOTES

1. Strictly speaking, a panel is more than that; in other words, it is possible to create a "pool" of cross-sections, which is not a panel per se.
2. When we talk about micro or macro data sets, we are describing the source of the observations, not the type of analysis the data are used for. One can test a macro hypothesis (e.g., a consumption function) with micro data, and vice versa (e.g., a demand function for an aggregate consumption good or service, using the decomposition of quantity and price data for personal consumption expenditures).
3. For nations other than the United States, there are comparable government statistical agencies. See, for example, the country entries under "World and Non-U.S. Data" at *Resources for Economists on the Internet* (<http://www.rfe.org/Data/World/index.html>). A similar listing is available at (<http://labour.ceps.lu/statisticsframe.cfm>).
4. Data are also available for Washington, D.C.; Puerto Rico; and the U.S. Virgin Islands.

SUGGESTIONS FOR FURTHER READING

- Best (2001)*—Very thoughtful monograph on how data series are constructed, and the ways they can distort what they purport to measure.
- Brown et al. (1996)*—Very readable introduction to the Panel Study of Income Dynamics. Well worth reviewing if you are considering using the PSID.
- Clayton and Giesbrecht (2001)*—Thoughtful handbook on how the major macroeconomic statistics are computed, what they measure, and how to correctly use them. Discusses statistics on output, production and growth; investment and capital expenditures; employment, earnings, and profits; spending, sales, and expectations; the price level, money, and interest rates; and financial markets, international trade, and foreign exchange rates.
- Huff (1954)*—Classic monograph on the potential problems with commonly reported statistics. Very readable and intuitive.
- Maier (1999)*—Excellent guide to the major social science statistics. Includes information on how to obtain the data online, as well as commentary on how the data are obtained and on their reliability.
- Morgenstern (1979)*—Classic critique of the use of macroeconomic statistics. Anyone using macro data should be familiar with this. Reprinted from Morgenstern's *On the Accuracy of Economic Observations*, 1963.
- Pergamit et al. (2001)*—Helpful introduction to the National Longitudinal Surveys. Worth checking out if you are considering using the NLS.
- Popkin (2000)*—Decent introduction to the National Income and Product Accounts. Worth reading if you are considering using the NIPA.
- Teplin (2001)*—Very good introduction to the Flow of Funds Accounts. Well worth reading if you are considering using the Flow of Funds.

EXERCISES

1. Go to the FRED II database at the St. Louis Federal Reserve Bank online at <http://research.stlouisfed.org/fred2/>. Download the monthly unemployment rates for the past five years for the states in the St. Louis Federal Reserve region. Record the complete URL for each data series. Compare and contrast the state unemployment rates.
2. Go to Economagic.com. Collect data for 2000 per capita disposable personal income for your home state and all the contiguous states. Rank them from highest to lowest. Report the complete URLs.

3. Find the primary source for data gross state product. Collect the data for 2002 for any ten states in the United States. Report the complete URLs.
4. Suppose you need annual data for GDP for Germany, France, and the United Kingdom. What source would you go to and why would that be the best?
5. Suppose you are asked to obtain annual data for an aggregate production function for the United States: $Q = f[L, K]$. Track down sources for these three variables and report the complete URLs. Explain why you chose each data source.

Overview of Data Sources

Places to Start

Resources for Economists on the Internet (<http://rfe.org>)

WebEc (<http://netec.wustl.edu/WebEc>)

U.S. Macroeconomic Data

National Income and Product Accounts—Econmagic, FRED, STAT-USA

U.S. Flow of Funds Accounts (macro/financial data)

Other Nations' Macroeconomic Data

National Statistical Agencies for Selected Countries (<http://rfe.org/data/world/index.htm>)

EC Member Countries—EuroStat

OECD Member Countries—OECD Main Economic Indicators and OECD National Accounts

IMF Member Countries—IMF International Financial Statistics

Other Nations—World Economic Outlook, Penn World Tables, U.N. System of Accounts

U.S. Labor Market Statistics

Current Population Survey

Current Employment Statistics

Econmagic

U.S. Microeconomic Data

Agricultural products—U.S. Department of Agriculture

Production of Manufactured Goods—Survey/Census of Manufactures

Demand for Consumer Goods and Services—Survey of Personal Consumption Expenditures (NIPA Table 2.4—PCE by Type of Expenditure), or Consumer Expenditure Survey

Demand for individual companies' products—Hoover's Online and Edgar

U.S. Industry Data

The Economic Census

Annual Survey of Manufactures

Current Industrial Reports

U.S. Micro Data Sets

National Longitudinal Surveys

Panel Study of Income Dynamics

IPUMS

International/Trade Data

U.S. Balance of Payments Accounts

U.S. imports and exports—Bureau of Economic Analysis, Customs Bureau

IMF International Financial Statistics.

Eurostat

International Investment Position of the U.S.

World Bank Statistics on External Debt

chapter

9

Putting Together Your Data Set

"Economists . . . may have a tendency to place more faith in the accuracy of data than is sometimes warranted. . . . We should be diligent to remind ourselves that the mere existence of data in numerical form does not, in itself, make it accurate or error-free."

DON ETHERIDGE

In Chapter 8 we observed that data collection is a major part of any empirical research project. Empirical research can be divided into two types: experimental and survey (or nonexperimental). In experimental research, the data come from performing the experiment. Thus, in experimental research collecting the data is, by definition, a significant part of the project. Unfortunately, in survey research where one uses preexisting data, data collection is not always given the same respect; researchers often don't put the same care and effort into it. This is a serious mistake! If you think I'm wrong here, ask yourself the following: If you use the most advanced statistical techniques but your data are flawed, what kind of inferences can you draw from the results?

In this chapter we will explain what you need to do at a minimum to construct a satisfactory data set. We will discuss how to develop a search strategy for finding appropriate data with a minimum of time and effort, using both primary and secondary data sources. We will also discuss how to convert the data into a useable form and to construct a data appendix for your research.

Developing a Search Strategy for Finding Your Data

When you search for data, just as when you do your literature survey, it's a good idea to start with a search strategy. Time spent up front in this activity will almost always pay off by significantly reducing the effort required to find and collect the data you need. The process for developing a search strategy is outlined in Figure 9.1.

Step 1: Before You Search

At least three issues are involved in the construction of a good data set. The first issue is sample size—you need to have enough data points to derive statistically valid empirical test results. You typically need thirty data observations to ensure adequate degrees of freedom, though you can sometimes make do with less. The second issue is that of a random or representative sample. We touched on these two issues in the previous chapter, but we will discuss them in detail in Chapter 10.

The third issue is obtaining data that correctly measure the concepts that your theory deems important. Begin by reviewing your conceptual or theoretical model. What variables do you need data for to test your research hypothesis? If you could obtain the ideal data, what variables would they include? At the highest level of abstraction, these questions should be straightforward to answer. If not, you may need to rethink your research hypothesis to obtain more focus.

Next, come a series of practical questions: What are the different forms in which those data might be available? There are several aspects to this

Figure 9.1 Developing a Search Strategy for Finding Your Data

Step 1: Before You Start Searching

- What are the desired variables?
- How should each variable be defined?
- What data frequency and sample period or what level of analysis?
- What are potential sources for data on each variable?

Step 2: As You Search

- What data are available?
- Are there suitable proxy variables for variables that are unavailable?
- If not, how can the empirical model be modified to use the data available but still test the hypothesis?

query. The first is conceptual. For example, if your theory requires a measure of income, you might find two types: per capita income and family income. Which is the best measure to test your hypothesis ideally?

The second aspect to this question is whether you should seek time-series or cross-section data. If the former, what frequency of data should you look for: annual, quarterly, monthly, weekly, or daily, and over what sample period? If the latter, what unit of analysis is most appropriate for your hypothesis: average income at the state level or the county-level, or even individual family incomes?

Since there may be a variety of ways to measure the variable you are interested in, you should think carefully about what your variable would be ideally, and then what you can live with.

Once you have determined your list of desired variables, the next step is to think about where those data are likely to be found—for example, by using the sources discussed in Chapter 8. What agency collects that data? What data sources were used by previous studies on your topic?

A good place to start your data search is the website, Resources for Economists on the Internet (RFE, <http://rfe.org>). RFE is an extensive source of a variety of information useful for economists. You may recall that we introduced it in Chapter 3 as part of our discussion on surveying the literature. RFE provides an extensive and annotated list of data sites (including a number of more specialized sources omitted from our discussion in Chapter 8). The annotations are particularly useful since they help you determine if a data site will be useful for your research, without your having to dig into it yourself. RFE's data collection is divided into five sections: U.S. macro and regional data, other U.S. data (e.g., data for microeconomic research), world and non-U.S. data, data from financial markets, and data from published journal articles.

Another useful resource for beginning a data search is the website, World Wide Web Resources in Economics or WebEc (<http://netec.wustl.edu/WebEc>). This site provides information similar to that of RFE, but with more of an international focus.

Once you have reviewed RFE and/or WebEc, you should have some good ideas about where to look for the data you need. These are likely to include both primary and secondary data sources. You may wish to review Appendix 8A, which outlined the primary and secondary data sources described in Chapter 8.

Step 2: As You Search

If you have completed Step 1, you now have a list of ideal variables for testing your hypothesis. You have also identified several promising data sources, such as Economagic and FRED II. The next step is to begin the search.

As you investigate each data source, you need to ask several questions:

1. **What data are, in fact, available?** Remember that this question involves both correctly defined variables and a sufficient number of data points.
2. **If the data are not the ideal, are they good enough?** In practice, it is uncommon for the ideal data to exist with a sufficient number of observations. After all, the folks who produced the data probably didn't have your exact research in mind. You might find the exact variable you desire but too few data points. Or you might find sufficient data points but for a variable that is not defined exactly as you'd like.
3. **If the data are not acceptable, is there an available proxy that is?** A proxy is a variable that should behave roughly the same as your theoretical variable. For example, it should have the same trends and turning points as the theoretical variable, though it isn't conceptually identical.
4. **If there is no adequate proxy, how can the hypothesis be reformulated to make it testable, given the data available?**

Let's illustrate these issues with a series of examples. Suppose you would like to test the Keynesian consumption function. The theory hypothesizes that consumer spending in the aggregate is a linear function of national income or GDP.

$$C = a + b * GDP$$

If you wanted to test this theory for the U.S. economy using annual or quarterly data, you would have no difficulty obtaining a large number of observations for both C and GDP. Macro data for the U.S. economy are among the best data in the world; they are available in great detail over long time frames.

But suppose you wanted to apply this theory to a foreign country. Such data are often more difficult to obtain, if they are obtainable at all. One complication, especially when doing international comparisons, is that concepts may be measured differently in different countries. For example, the unemployment rate in Japan is not defined in the same way as in U.S. data. Additionally, in many countries published macroeconomic data are thought to be politically manipulated. Thus, the published data, even if available, may not correctly measure the underlying theoretical concepts.

Consider another example closer to home. Suppose you wished to test the Keynesian consumption function over a monthly time frame. Data on personal consumption expenditure and GDP are available in annual or

quarterly frequency, but not monthly. Thus, you would need to find proxy variables in order to do a monthly study. Recall what makes a good proxy. Either the proxy should measure the same underlying economic behavior as the theoretical variable you wish to study, or it should be highly correlated in a statistical sense with the theoretical variable. For example, retail sales are published monthly, and they measure approximately the same thing as consumption expenditure. Finding an adequate proxy for GDP is more problematic. One candidate might be the Fed's index of industrial production, which is available monthly. Industry production measures the manufacturing output of the economy. In the early twentieth century, when the U.S. economy was largely industrial, this would have made a good proxy for GDP. In recent years, however, manufacturing has declined to less than 40 percent of GDP, while the service sector is roughly 60 percent. To see if industrial production is an acceptable statistical proxy, one should examine the correlation between the two variables over the closest frequency available, in this case, quarterly. Examining data for quarterly real GDP and quarterly industrial production indicates a very high correlation of 0.96. Thus, industrial production probably is an acceptable proxy.¹

These problems are often more acute when you are doing a microeconomic study. For these, you will more likely need to use variables from different sources, and perhaps different definitions. It is also more likely that different variables may come from different populations. Suppose your theory identifies weekly work hours by couples with children between the ages of six and seventeen. The available data may include weekly work hours by gender nationwide as well as employment rates among males and females in couples with children. You can construct a data series that approximates the theoretical concept, by making suitable assumptions. For this example, you could assume that males and females with children have, if employed, the same weekly work hours as males and females in the general population. You still need to be concerned about how well the resulting series serves as a proxy for the desired theoretical variable. Ultimately, to perform the analysis at all, you may have to hope for the best and use the data, but acknowledge its potential weaknesses.

Oftentimes, researchers must simplify their theoretical model because they lack adequate data on the variables they need. Either there are insufficient data points for one or more variables, or the specific variable they need is not available. In our example of the monthly consumption function, if a proxy for monthly GDP cannot be found, you may need to change the study to examine a quarterly consumption function.

Figure 9.2 Different Forms of Data

- Levels
- Per capita
- Changes
- Rates of Change (or Growth Rates)
- Annualized Growth Rates
- Proportions
- Nominal
- Real
- Index Numbers

Data Manipulation

Data for any variable are available in a number of different forms. The most important of these are listed in Figure 9.2. In general, it is possible to convert a variable from one form into another. In this section we summarize some commonly used techniques for performing those conversions. Even if you can find data in the exact form you need so you don't need to perform these calculations, it is nonetheless important that you understand what each form means and how they relate to each other.

Level of Variable

The most basic form of any variable is called the **level**. The level is the actual value or size of the variable being measured. For example, Figure 9.3 shows that the level of U.S. GDP in 2001 was \$10,082 billion.

Often, researchers prefer to use the **per capita** form of a variable. Per capita means per person. It is found by dividing the level of the variable by the appropriate population. For example, since the U.S. population in 2001 was 285.318 million, U.S. GDP per capita in 2001 was \$10,082

Figure 9.3 Levels, Changes, and Growth Rates of U.S. GDP (in billions of \$)

Year	GDP	Change in GDP	Percentage Change
1999	\$9,274.3		
2000	\$9,824.6	\$550.3	5.9%
2001	\$10,082.2	\$257.6	2.6%
2002	\$10,446.2	\$364.0	3.6%

NOTES FOR NOVICE RESEARCHERS

Pay Attention to the Units!

Novice researchers often pay little or no attention to the units in which variables are measured. This is a mistake you should avoid! In our previous example of GDP per capita, GDP is measured in billions of dollars, while the population is measured in millions of people. Dividing the two yields units of thousands of dollars per person. Note that it's not enough to know that GDP is measured in dollars—you should know whether it's trillions of dollars or millions of dollars. Suppose you ignored those zeros in GDP and population, and only divided the digits: 10,082 / 285.318. You would have obtained an answer of \$35.37 per capita. So if you think that the number of zeroes in statistics doesn't matter, ask yourself whether you'd prefer to have an annual income of \$35,000 or \$35.

billion / 285.318 million = \$10,082,000 million / 285.318 million = \$35,336 per person.

One common use of the per capita form of GDP is to compare standards of living between nations. Mexico and Spain have approximately the same size economy: Mexico's 2002 GDP was about \$900 billion, while Spain's was \$828 billion, so Mexico's was slightly larger. However, Mexico's population is substantially larger than Spain's, so Spain's standard of living is much higher. Mexico's 2002 GDP per capita is about \$9,000 per person; by contrast, Spain's 2002 GDP per capita is more than double that, at about \$20,700 per person. For similar reasons, when you are performing other types of comparisons, you might use the per family or per household form of income or some other variable.

Change in Variable

For some purposes, researchers find it more useful to examine the **change** in a variable than the level. For example, knowing that the level of U.S. GDP in 2001 was \$10,082 billion may be less important than knowing that it increased by \$258 billion between 2000 and 2001. In other words, the change in a variable is just the difference in the levels from one time period to another.

Often, it is important to know whether a change was meaningful. Was the increase in GDP between 2000 and 2001 large or small? One way to assess

NOTES FOR NOVICE RESEARCHERS

Caution on International Comparisons

Each country measures its economic data, such as GDP, in its own national currency. In other words, Mexico's GDP is measured in pesos, while Spain's GDP is measured in pesetas. In order to make cross-country comparisons, as we did earlier, the data need to be converted into a common unit of account, a common denominator. Either both GDPs need to be in pesos, both in pesetas, or both in some third currency, like dollars. This conversion can be done in two ways. The most common way is to use an exchange rate, that is, the price of one currency expressed in terms of another. Thus, one could take Spain's GDP (in pesetas) and multiply by the U.S.\$-peseta exchange rate to obtain Spain's GDP in terms of dollars. Going through the same process for Mexico, we would multiply its GDP in pesos by the U.S.\$-peso exchange rate to yield Mexican GDP in terms of dollars. Then the two GDPs could be compared.

In practice, however, this is not the best way to convert Mexican and Spanish GDP. This is because exchange rates fluctuate regularly, sometimes dramatically in a short period of time, which makes it appear that a country's GDP has increased or decreased when really it has not. For many purposes, then, such as comparing two nations' standards of living, it is preferable to use *purchasing power parity equivalents* (or PPP Equivalent) instead of exchange rates to convert the GDP figures. Indeed, the figures we just cited were computed with this method. The PPP Equivalent is a measure of the relative purchasing power of each country's currency.² Essentially, it extracts the volatility from exchange rates to yield a more accurate, long-term measure of the value of a currency in terms of another. Whenever possible, you should use the purchasing power equivalent version of economic data when comparing across countries. This is usually labeled "purchasing power parity \$."

Note also that direct comparisons of GDP or income per capita can be misleading when you are making comparisons between economies that have different economic systems or levels of economic development. For example, the reported income per capita of rural residents in a country tends to undervalue their actual standard of living, since they often have unreported income-in-kind such as

produce they grow for their personal consumption. In other words, comparing reported incomes per capita between urban and rural residents would be biased in favor of the former group. The same problem occurs when comparing incomes per capita between developed and developing nations.

this is to look at the **rate of change**, also called the **percentage change** or the **growth rate**. The rate of change is the change as a percentage of the original level. To compute the rate of change for some variable X between an initial period 0 and a subsequent period 1, you can use the following formula:

$$G_{1,0} = \{[X_1/X_0] - 1\} * 100$$

where X_0 is the level (or value) of X in the initial period, X_1 is the level of X in the subsequent period, and $G_{1,0}$ is the rate of change or growth rate between periods 0 and 1.

For example, the growth rate of GDP from 2000 to 2001 was

$$\begin{aligned} & \{[\$10,082.2 \text{ billion}/\$9,824.6 \text{ billion}] - 1\} * 100 \\ &= \{[1.0262] - 1\} * 100 \\ &= \{.0262\} * 100 \\ &= 2.62\% \end{aligned}$$

Growth rates are especially useful for comparing concepts that are different orders of magnitude, like the annual sales of Microsoft versus the annual sales of Joe's Computer Software, Inc.

By convention, researchers usually compare data in terms of annual growth rates, even if the period of time they are examining is more or less than a year. The reason for this is so they have a common standard for comparison. Growth of GDP over a decade is almost certainly greater than growth over a shorter period, but it isn't clear that GDP is growing at a faster or slower rate over the decade until you standardize the two time periods examined. For periods longer than a year, the overall growth rate is simply divided by the number of years involved. Thus, if GDP grew 20 percent over ten years, the average annual growth rate would be $20/10 = 2\%$ per year.

For periods of time shorter than a year, like a month or a quarter, researchers usually follow a different procedure. They ask the question, "If growth was X% for some period less than a year, what would the cumulative rate be if that growth rate continued for an entire year?" This is called an **annualized growth rate**. You might think that you could just multiply

Figure 9.4 Annualized Growth Rates

Beginning:	Quarter I	Quarter II	Quarter III	Quarter IV
	= 1.10*100	= 1.10*110	= 1.10*121	= 1.10*133.1
Sales:				
\$100	\$110	\$121	\$133.1	\$146.4
				46.4%

quarterly data by four or monthly data by twelve to get annual rates. That would in fact provide a rough approximation, but only a rough one because it ignores the effects of compounding. (You may recall the way that interest on a bank account compounds. Not only do you earn interest on your deposit, but also you earn interest on the interest. That's compounding.)

In the same way, suppose that a company's sales were growing at 10 percent during the first quarter of the year. What would that amount to at an annual rate, assuming that the 10 percent growth rate continued for three more quarters? Figure 9.4 illustrates this question. Suppose the level of sales was \$100 at the beginning of the year. At the end of the first quarter, sales would be 10 percent higher or \$110. At the end of the second quarter, sales would be 10 percent higher than that, or \$121. By the end of the fourth quarter, sales would be 46.4 percent higher than when the year started. In this case, since growth compounded over four quarters, the annualized growth rate (46.4%) is 6.4 percent more than the rough approximation ($4 * 10\% = 40\%$). If you were using monthly data instead of quarterly, the annualized growth rate would be even more than the approximation, since the compounding occurs twelve times. In other words, the more frequently the compounding is calculated, the greater the error in the approximation.

To compute an annualized growth rate for quarterly data, use the following formula:

$$G_q = \{[X_1/X_0]^4 - 1\} * 100$$

where X_0 is the level of X in the initial quarter, X_1 is the level of X in the next quarter, and G_q is the annualized rate of change or growth rate. The formula for annualizing monthly data is the same except that instead of quarterly values of X you would use monthly values. In addition, instead of raising the ratio of quarterly values to the fourth power, you would raise the ratio of monthly values to the twelfth power:

$$G_q = \{[X_1/X_0]^{12} - 1\} * 100$$

Figure 9.5 Proportions of 2002 U.S. GDP

(Values for GDP and components are in billions of \$)

GDP	C	I	G	X-M
\$10446.2	\$7303.7	\$1593.2	\$1972.9	-\$423.6
	= 7303.7/ 10446.2	= 1593.2/ 10446.2	= 1972.9/ 10446.2	= 423.6/ 10446.2
	69.9%	15.3%	18.9%	-4.1%

A form of data similar to a growth rate is a **proportion**. A proportion is also called a share or a **percentage** or a fraction. The idea behind proportions is simple: imagine dividing the economic pie into several parts. Each part is a proportion. Thus, labor income (i.e., wages, salaries, etc.) is roughly 75 percent of national income. Similarly, capital income (i.e., interest, rentals, etc.) is roughly 25 percent of national income. A proportion is measured as a fraction or percentage of the whole. Consider the example in Figure 9.5. Any student of macroeconomics knows that GDP can be divided into consumption expenditure, investment expenditure, government expenditure, and net export expenditure. Each of these fractions is a proportion. The proportion of GDP that is consumption expenditure is computed as the level of consumption expenditure divided by the level of GDP, or 69.9 percent.

Proportions are useful for understanding a wide range of issues in economics. For example, Americans are thought to have a low saving rate. The saving rate is just the proportion of income that is saved. Is the amount of poverty in America growing or shrinking? One way to assess that is by looking at the poverty rate, which is the proportion of Americans with income below the poverty line.

Real versus Nominal Magnitudes

A great deal of the available economic data are based on a simple identity:

$$V = P \times Q$$

where V = nominal (or value)

P = price, and

Q = real (or volume).

Nominal data are data measured by using the actual market prices that existed during the time period in question. For example, the value of a firm's output in 1999 is equal to the product of the price the firm charged

in 1999 and the quantity of output it produced in that year. Similarly, nominal GDP in 1999 is the sum of the price times the quantity of every product produced in a nation in 1999. Nominal data are easy to interpret.

Real data are a bit more difficult to understand. At the micro level, real data refer to the actual quantities employed by a firm (e.g., labor hours), produced by a firm (e.g., number of widgets), or sold by a firm (i.e., sales volume). In our example, the firm's real output is the actual quantity of goods or services that it produced.

The story becomes more complex, however, when we move above the level of an individual firm to a more macro perspective. How does one measure real GDP, the actual amounts of goods and services produced by a nation? One could create a list of all the products produced in every industry. But aside from being cumbersome, such a real measure of GDP would make it difficult or impossible to compare GDP from one year to the next. If last year a nation produced 50,000 automobiles and 17,000 new homes, while this year the figures were 52,000 and 14,000, did GDP rise or fall? The solution economists have found to deal with this quandary is to define real GDP as a weighted average of all the goods and services produced, where the weights employed are the prices of each product in a given year.³ In other words, money provides the common denominator by which to measure diverse goods and services produced in a year.⁴

Index Numbers

Real GDP, as just defined, is an example of what economists call an **index number**, or more specifically, a **quantity index**, since it measures the average quantity of something. The other type of index number is a **price index**. These are more familiar to noneconomists. For example, most people have heard of the consumer price index. What is a price index? A price index is (roughly speaking) a weighted average of the prices of some type of good or service, where the weights are the corresponding quantities in a given year.⁵ One could construct an energy price index (the average price of different energy sources), a food price index (the average price of different foods), or a consumer price index (the average price of goods and services consumers typically buy). An index number is a data series that begins at 100 and typically increases in subsequent years, though not necessarily continuously. For example, Figure 9.6 gives the consumer price index for 1984–1993.

How does one interpret an index number? Suppose you learn that in July 2003 the consumer price index had a value of 183.9. What exactly does that mean? The short answer is that 183.9 doesn't mean anything in isolation. Index numbers, unlike most other statistics, have no units. Rather, index numbers are designed for comparison purposes. For example, one

Figure 9.6 Consumer Price Index, 1984–1993

1984	1985	1986	1987	1988	1989	1990	1991	1992	1993
103.9	107.6	109.6	113.6	118.3	124.0	130.7	136.2	140.4	144.0

could use an index number to compare the level of whatever the index is measuring (for example, consumer prices) to some earlier time period known as the base period. Thus, a CPI of 183.9 means that on average consumer prices have risen 83.9 percent since the early 1980s, which was the base period for these figures.

Index numbers are typically constructed by taking the raw index, that is, the actual weighted average just mentioned, and transforming it using the following formula:

$$X_T = [X_t / X_0] * 100$$

The variables in the formula are defined as follows:

- X_t is the value of the raw variable in a given time period t in the series,
- X_0 is the value of the raw variable in the base period, which is the period to be compared against, and
- X_T is the resulting index number.

Note that in the base period $X_t = X_0$. Therefore, an index number always has a value of 100 in the base period.

Index numbers can also be used to compare the relative growth rates of two similar concepts, such as employment in Colorado and Virginia. They are particularly useful when the concepts have very different magnitudes. Imagine trying to graph employment in your local town and employment in the nation as a whole over the past ten years, both on the same graph. The national figures would be so much larger than the local ones as to make it very difficult to compare them.

To make it easier to see what was going on, you could index both data series to a common starting point, say, ten years ago. What you would be doing is transforming each series to have an arbitrary value of 100 in that year, using the formula for computing index numbers given earlier. Then since each would have a common starting point, you could easily see how employment growth in your town compared to the national economy.

Quantity Indices Versus Real Quantities

There are two ways to construct real measurements. The first is to create a quantity index, as described earlier (i.e., a weighted average of the quantities

of the type of good or service one is interested in). The second is to create a price index and then divide the nominal value by the price index. The result is real and follows directly from the identity above: $Q = V/P$. The data for real GDP in the U.S. National Income and Product Accounts are created using the second approach.

Price Indices Versus Implicit Price Deflators

There are similarly two ways to construct price measurements. The first is to create a price index as described earlier (i.e., a weighted average of the prices of the type of good or service one is interested in). The second is to create a quantity index and then divide the nominal value by the quantity index. The result is called an **implicit price deflator**, since it is the price index implied by the identity, which is expressed in this case as $P = V/Q$. Note that published price indices are multiplied by one hundred.

You might think that a price index and the corresponding implicit price deflator would be the same. However, though we tend to use them equivalently, only under extreme technical assumptions will that be the case.⁶

How Inflation Distorts Nominal Values

Because prices tend to increase over time, comparing nominal measurements can be misleading. For example, if you receive a 10 percent salary increase from one year to next, while at the same time the prices of the things you typically buy experience 10 percent inflation, you will not enjoy an increase in your standard of living equivalent to your nominal raise. In fact, your living standard, or more precisely your real income, will not have changed at all. This is the concept behind another way to describe real measures: as nominal measurements adjusted (or corrected) for the effects of inflation.

The way this “inflation adjustment” is done is by recalling that GDP is a weighted average of Q , where the P s are prices in a specific period, called the base year. By using base-year prices but actual-year quantities, real GDP excludes the effects of changing prices over time. (Note that if relative prices change significantly, that can cause a bias in the measure.)

Rebasing Data Series

Periodically, those who collect the data change the base year. One reason they do this involves how one interprets the base year from the point of view of the index number. That is, the base year is the period against which the index measures changes in prices or quantities. We care more about recent economic changes than historic ones, so the base year is changed to keep it “recent.” A second reason for changing the base year is to correct the bias mentioned in the last section.

NOTES FOR NOVICE RESEARCHERS

Practical Data Tip

Any nominal variable can be transformed into its corresponding real magnitude by dividing by an appropriate price index: $Q = V/P$. An easy way to see this is by examining Tables B-1, B-2, and B-3 of the *Economic Report of the President*, which show nominal GDP, real GDP, and the GDP deflator, respectively.⁷

Similarly, if one has data for nominal and real (for example, export sales and export volumes), one can derive a price index: $P = V/Q$. Note that such an index is called a **unit value**.

Re-basing a data series can cause difficulties for researchers. This is particularly a problem with print data sources. When you compile data from print sources, you should always start with the most recent source and work backward. This will enable you to capture revisions in the data that were not available until a subsequent date. For example, suppose you are compiling an annual data series, and each year your source reports the estimated figure for the current year, and also revised data for the previous two years. If you start from the earliest source and go forward, your data series will consist of the earliest and least reliable estimates of the variable.

Suppose you have a time series of annual CPI data in which the first half of the data uses a base year of 1992, while the second half uses a base year of 1997. There is no perfect way to link the two series! Every way imposes some bias, but that is the nature of working with real-world data. To link the two parts of the series it is necessary to have some overlap, as shown in Figure 9.7 for the year 1995. The first row shows the source data with a base year of 1997. (This should be clear since the 1997 value for the data is 100.) The second row shows the source data with a base year of 1992. Notice in particular that if we merely combined the two series there would be a noticeable jump where the two series connect. In short, the data would show an incorrect change in the CPI in that time period. For example, if you combined the base 1992 data through 1994 with the base 1997 data beginning in 1995, you would see a large apparent decrease in the CPI between 1994 (i.e., 1.30) and 1995 (i.e., 0.85). This would indicate deflation, when in fact prices are rising!

Figure 9.7 Rebasing a Data Series

Year:	1991	1992	1993	1994	1995	1996	1997	1998
Base Year 1997					0.85	0.95	1.00	1.15
Base Year 1992	0.95	1.00	1.25	1.30	1.40			
Linked Series	0.58	0.61	0.76	0.79	0.85	0.95	1.00	1.15

Suppose we want a complete series with a 1997 base year. We will need to transform the values for the observations that are only available with a base year of 1992 so they correctly show the change in CPI between both parts of the source data. The data with the earlier base year need to be reduced in value to correctly link to the data with the later base year. The amount of the reduction is given by the ratio of the two values for 1995, the period of overlap. In other words, dividing the base-year 1997 value into the base-year 1992 value yields the reduction factor: $85/140 = 0.607$.

To obtain the linked series with a 1997 base year, use this factor to multiply the value of each observation with base-year 1992. The result is shown in the third row of Figure 9.7.⁸ Notice that there is no longer a drop in the series where the two parts were linked.

Data Smoothing

Data that are volatile are sometimes “smoothed” to better reveal the underlying trends. Consider the monthly income of a salesperson who works entirely on commission. One month his income might be quite high, while the next it could be low. That variability makes it difficult for the salesperson to tell if he’s having a good year or not. The problem is similar for researchers using certain types of data.

There are a variety of techniques for smoothing data. We will discuss two such techniques here: moving averages and seasonal adjustment. A **moving average** replaces the actual data point in each period with an average of the $n - 1$ preceding data points with the n th. The result is that any “abnormal” observations become less important since they are averaged with more normal ones.

Some variables show seasonal patterns, whereby they change predictably in certain months, quarters, or seasons. For example, consumer spending rises in the fourth quarter of every year in anticipation of Christ-

mas. Often, researchers wish to know if the changes are abnormal. Thus, one might ask, did consumers spend more or less than normal this Christmas season? This is a question for which **seasonally adjusted** data are designed. The U.S. Census Bureau has developed techniques that are commonly used to seasonally adjust data. The most recent version is called X12. This version of the earlier X11 technique is widely available on common statistical packages or it can be obtained free from the Census Bureau. Note that many macro data series are only available in seasonally adjusted form. These include GDP, CPI, PPI, retail sales, and housing starts. It is unlikely that you will need to do your own seasonal adjustment of data, but it is important that you understand what the term means. Note the following abbreviations, which are often used with data sources:

NSA: Not seasonally adjusted.

SA: Seasonally adjusted.

SAAR: Seasonally adjusted annual rate.

Appendix 11A discusses some additional ways to manipulate data. Whatever data manipulation you do, once you have your completed data set it is good practice to review the data. Examine each data series by looking at either data tables or charts. Look for observations that appear to be errors. This is easiest with time-series data that show a trend. A good example is the CPI, as shown earlier. Just because data is published does not mean that is error free. Moreover, if you performed any data entry, there is always the possibility you made a keystroke error. For cross-section data you should compute summary descriptive statistics: mean, range, and standard deviation for each variable. You can report these as part of your data appendix. This is not generally done for time-series data, which often contain a time trend. Since the purpose of the descriptive statistics is to give you a rough sense of what the data are like, with time-series data one would be likely to discover that the lowest data point was the earliest, while the highest data point was the latest. With a time trend, you already know that.

Only when you have examined the data and satisfied yourself that they are as accurate as they can be are you ready to begin statistical testing.

Constructing a Data Appendix to Your Research

It is good scientific practice to make your data available so that other researchers may replicate your work. At a minimum, this means explaining the “Sources and Methods” you used in compiling your data set. We can categorize data as **raw (or source) data** or **derived data**. Raw data are the data as you found it at the source. Derived data are the data that you actually used in your empirical testing. This distinction is important because it is not uncommon for the raw data you obtain to be in a form different

from what you need to test your hypothesis. For example, the data available from the source might be the consumer price index, while the data you need might be the inflation rate, that is, the percentage change in CPI. Thus, you may need to manipulate or mathematically transform the data before using it. In this example, the CPI would be the raw data, while the inflation rate would be the derived data.

In explaining your sources and methods you should provide clear citations for exactly which sources provided the raw data, as well as a complete explanation of how you manipulated the raw data to put it in the form you actually used. After reading your documentation, the reader should be able to replicate your exact data with a minimum of digging. Better yet, you should provide a copy of the exact data you used, in addition to documenting where it came from. A number of journals require that authors include copies of their data when they submit their research papers for publication.⁹ This copy of your data with documentation is called a data appendix.

SUMMARY

- Since the quality of your research will be no better than the quality of your data set, you should put a substantial effort into constructing a good quality data set.
- To efficiently locate data you should begin with a search strategy. Begin by creating a list of the desired variables for your research. Next, consider likely sources for that data. Think about what proxies might be available for data that are unavailable. Finally, consider how the model might be revised to use the available data.
- Data are available in different forms, including levels, per capita, changes, rates of change, and proportions.
- Growth rates for periods of less than a year are often given in annualized form.
- Nominal data are distorted by inflation; real data are corrected for inflation.
- An index number is essentially a weighted average of whatever the number purports to measure. Thus, the consumer price index (CPI) is a weighted average of the prices of goods and services consumers typically buy. Index numbers are designed to show changes in prices or quantities since some arbitrary base period.
- A data appendix provides the exact data the researcher used as well as a complete description of the sources and methods he or she used to obtain it.

NOTES

1. Monthly values of industrial production were averaged to obtain a quarterly series. Real GDP is available quarterly. Correlation will be explained in detail in Chapter 10.
2. The purchasing power parity (PPP) equivalent gets its name from the fact that it measures currencies in terms of the goods they can buy. PPP is based on the notion that in equilibrium the same product should cost the same amount in two different currencies when adjusted by the exchange rate. This should be true in equilibrium; if these costs are not the same, it would be profitable to export from where it is cheaper and import to where it is more expensive until the costs are equal. Therefore, PPP is a long-term equilibrium measure of the exchange rate.
3. Thoughtful readers may notice that this is the same value as nominal GDP. This point will be clarified shortly.
4. This is exactly what is meant by characterizing money as “a unit of account.”
5. For some price indices the weights are budget shares rather than quantities; in other words, the weight of some consumer product price is the proportion of that item in the average consumer’s budget.
6. For similar reasons, a quantity index and the corresponding real quantity are rarely equal, though this is easier to see because the units are different. Real quantities are measured in constant (e.g., 1992 dollars), while quantity indices like all index numbers are unit free, that is, they are numbers like 100, 117, 129. This will be explained in more detail later in this chapter.
7. Note, however, that for the math to work out correctly the data for the GDP deflator must be divided by 100.
8. The careful reader will notice that this procedure for rebasing index numbers is a variation on the formula for creating index numbers presented earlier.
9. Recall that the last section of data in the Resources for Economists website was data from published journal articles.

SUGGESTIONS FOR FURTHER READING

DataBasics, Federal Reserve Bank of Dallas website, <http://dallasfed.org/data/basics/index.html>—A concise summary of the major techniques for manipulating data.

Wyrick (1994)—Offers explanations of a variety of data manipulation techniques.

EXERCISES

1. Suppose you are researching the question of why the U.S. unemployment rate for African Americans is higher than the U.S. average unemployment rate. Create a list of variables that you think might explain that question. For each variable, propose a source for that data. Next, visit each source, and describe the data you found. For example, if there are time-series data, what frequency is available and for what time period? If there are cross-section data, what level of analysis is available? Draw a conclusion as to whether the data is adequate to test the hypothesis empirically.
2. Go to the Bureau of Economic Analysis Fixed Assets webpage at <http://www.bea.doc.gov/bea/dn/faweb/AllFATables.asp#S1>. Download data for Current-Cost Net Stock of Fixed Assets and Consumer Durable Goods (Table 1.1) for the past ten years available. Open the downloaded file in a spreadsheet. Compute the year-to-year growth rate for nonresidential fixed assets. This is the capital stock of private business in the United States, in other words, K in the standard production function $Q = f[L, K, t]$. Is there a trend in the growth rates? Are they increasing, decreasing, or staying about the same?
3. Go to the Census Bureau's webpage on U.S. Trade Balances by Country at <http://www.census.gov/foreign-trade/balance/index.html>. What were U.S. imports from Mexico, Canada, Japan, and China in 2001? Next, go to <http://www.census.gov/foreign-trade/statistics/historical/gands.txt> to obtain total U.S. imports in 2001. Finally, compute the proportion of total U.S. imports in 2001 that came from Mexico, Canada, Japan, and China.
4. Collect data on quarterly nominal U.S. GDP for the past five years available. One source of this data is the Bureau of Economic Analysis's webpage at <http://www.bea.doc.gov/bea/dn/nipaweb/TableViewFixed.asp?SelectedTable=3&FirstYear=2002&LastYear=2003&Freq=Qtr>. Next, collect data for the U.S. implicit price deflator for GDP for the same time period, at the same webpage. Now, deflate the nominal GDP using the GDP deflator to get real GDP. Please show your work. Finally, to check your work, collect data for real GDP at the same BEA webpage. Compare your computed real GDP with the data you downloaded.
5. For each of the variables used in Exercise 1, provide complete citation information on how you obtained the data.

What Are Chained-Dollars?

In the chapter, we defined an index number as a weighted average. For example, a quantity index for real GDP would be a weighted average of the quantities of all products in GDP, where the weights are the prices of those products. To be more precise, most index numbers are constructed as the *ratio* of two weighted averages. In other words, a quantity index for real GDP would be the ratio of average production in the current period to average production in the base period. For a quantity index, the natural choice for weights is the corresponding prices, but which prices? Traditionally, there were two choices: the base period prices or the current period prices. If the former was chosen, the result was called a *Laspeyres index*; if the latter was chosen the result was called a *Paasche index*. For example, the consumer price index is computed using base period quantities as weights; thus, the CPI is a Laspeyres index. Each of these choices introduces a bias into the results. The bias of the Laspeyres index is in the opposite direction of the bias in the Paasche index. In recent years, a new weighting scheme has been introduced by the U.S. Bureau of Economic Analysis to deal with this issue. The result is called a *chain-type index*, which is defined as the geometric mean of the Paasche and Laspeyres indices. Note that unlike the fixed weights of a Laspeyres or Paasche index, for a chain-type index the weights change for every observation.

Example of a Data Appendix

Military Enlistment = $f(\text{Unemployment Rate, Military/Civilian Pay Ratio, War Status})$

Air Force, Army, Marines, Navy—Number of active duty enlisted soldiers in each of the respective branches of the U.S. military

U-Rate—The national Unemployment Rate averaged from monthly observations to yield annual rates

Pay Ratio—Ratio of military to civilian pay. Military Pay is measured as the monthly basic pay for active duty enlisted men and women of classification E-3 with two or fewer years of service. These monthly observations were divided by 160 to represent a 40-hour workweek. Civilian Pay is defined as the hourly earnings of civilian production workers. Monthly observations were averaged to produce annual averages.

War—This is a dummy variable with 1's for the years during which the United States had a major foreign conflict.

YEAR	AIR FORCE	ARMY	MARINES	NAVY	U-RATE	PAY RATIO	WAR
1981	570,302	781,419	190,620	540,219	7.616667	0.540545	0
1982	582,845	780,391	192,380	552,996	9.708333	0.531489	0
1983	592,044	779,643	194,089	557,573	9.6	0	1
1984	597,125	780,180	196,214	564,638	7.50833	0.512308	1
1985	601,515	780,787	198,025	570,705	7.191667	0.517612	0
1986	608,199	780,980	198,814	581,119	7	0.521581	0
1987	607,035	780,815	199,525	586,842	6.175	0.524918	0
1988	576,446	771,847	197,350	592,570	5.491667	0.518359	0
1989	570,880	769,741	196,956	592,652	5.258333	0.51926	1
1990	535,233	732,403	196,652	579,417	5.616667	0.517419	1
1991	510,432	710,821	194,040	570,262	6.85	0.522679	1
1992	470,315	610,450	184,529	541,886	7.491667	0.531482	1
1993	444,351	572,423	178,379	509,950	6.908333	0.537681	1
1994	426,327	541,343	174,158	468,662	6.1	0.535501	1
1995	400,409	508,559	174,639	434,617	5.591667	0.534311	0
1996	389,001	491,103	174,883	416,735	5.408333	0.529458	0
1997	377,385	491,707	173,906	395,564	4.941667	0.52527	0

1998	367,470	483,880	173,142	382,338	4.5	0.518798	0
1999	360,590	479,426	172,641	373,046	4.216667	0.518653	1
2000	355,654	482,170	173,321	373,193	3.975	0.522991	0
2001	353,571	480,801	172,934	377,810	4.758333	0.522497	1
2002	368,251	486,542	173,733	385,051	5.775	0.544941	1

Data Sources:

Bureau of Labor Statistics, *Data Series LNS 14000000: Seasonally Adjusted National Unemployment Rate*, http://data.bls.gov/PDQ/servlet/SurveyOutputServlet?data_tool=latest_numbers&series_id=LNS14000000 (accessed November 15, 2004).

Used for **U-Rate**

Bureau of Labor Statistics, *Data Series CEU0500000006: Average Hourly Earnings of Production Workers*, [http://data.bls.gov/PDQ/servlet/SurveyOutputServlet;jsessionid=f0307d7548c5\\$3F7\\$3Ft](http://data.bls.gov/PDQ/servlet/SurveyOutputServlet;jsessionid=f0307d7548c5$3F7$3Ft) (accessed November 15, 2004).

Used for **Pay Ratio**

Defense and Accounting Service, *Military Pay Prior Rates*, <http://www.dfas.mil/money/milpay/priorpay/> (accessed November 16, 2004).

Used for **Pay Ratio**

Department of Defense, *DoD Active Duty Military Personnel Strength Levels, Fiscal Years 1950–2002*, <http://www.dior.whs.mil/mmids/military/ms9.pdf> (accessed November 12, 2004).

Used for **Air Force, Army, Marines, Navy**

Military.com, 2004, *Military.com's Military History: The Complete List*, http://www.military.com/Resources/HistorySubmittedFileView?file=history_completelist.htm (accessed November 12, 2004).

Used for **War**

Source: Data appendix courtesy of David Hutchinson.

A First Look at Empirical Testing: Creating a Valid Research Design

"No amount of experimentation can ever prove me right; a single experiment can prove me wrong."

ALBERT EINSTEIN

Throughout this book we have argued that research involves creating a persuasive argument and that researchers persuade by providing theoretical and empirical evidence that supports their thesis. This chapter and the following one will explain how to develop convincing empirical evidence.

In this chapter, we introduce the concept of research design, or how to create a valid empirical test of a hypothesis. We survey three types of increasingly sophisticated empirical analysis: casual empiricism, simple statistical hypothesis testing, and multiple regression. Then we discuss common problems of data analysis that weaken or limit the validity of these empirical tests. We conclude by explaining why regression analysis is the empirical tool of choice for economists. In Chapter 11 we survey regression analysis in detail.

Key Issues of Research Design

Okay. You have an interesting research question. You've reviewed the literature on the topic area. You have analyzed the issue using economic theory, and from that theoretical analysis you have derived your hypothesis, that is, your proposed answer to the research question. Next, you need to test the hypothesis. In other words, you must assess whether or not your answer to the research question is correct. How do you design an appropriate test? There are a number of key issues to consider.

Two general types of empirical methodology are available to researchers: experimental and survey (or nonexperimental) methods. **Experimental** methods are illustrated by "laboratory" experiments. For example, the researcher might examine two sample groups, an intervention or treatment group, and a control group. The control group is designed to be in every way identical to the treatment group except for the treatment. Thus, if the experiment between the two groups yields different outcomes, then that difference can be ascribed to the treatment.

Suppose a professor teaches two sections of the same course. The two sections have the same lectures, assignments, and examinations, but in one section the instructor is trying a new textbook. If that section receives higher grades in the course, it could indicate that the new text was more effective than the other text. Of course, other possibilities need to be considered as well. It may be that students in the more successful section were brighter or more experienced in economics. Thus, in a well-designed experiment, the students would need to be randomly distributed between the two sections to control for outside factors.

Traditionally, research in economics has used **survey** or **nonexperimental** methods. (In recent years, however, experimental economics has been growing in importance). Survey research involves the passive observation and analysis of events as they occur in nature. For example, instead of the Federal Reserve inducing a recession, and then analyzing the outcomes, Fed economists use data on recessions as they have occurred in history to investigate their causes and effects. Survey research uses statistical methods to account for outside factors. This will be discussed later in this chapter in detail.

A critical factor in designing an empirical study is the degree to which the methodology is valid. What makes for a valid empirical research design? The **validity** of a study has multiple dimensions. At the broadest level, one needs to consider internal and external validity.

A study has **internal validity** if the impact observed can be attributed to the study variable, in other words, if given the assumptions and evidence one can deduce that X causes Y.¹ Returning to our teaching experiment, if the two class sections are identical except for the textbook used, we can logically conclude that the difference in grades is attributable to the choice of text. In short, the experiment has internal validity.

There are several elements to consider when assessing internal validity. The three most important are instrument validity, relationship validity, and causal validity.² **Instrument validity** asks whether the test instrument adequately measures what it purports to. For example, it is not uncommon for researchers to have difficulty finding ideal data. Are the available data

sufficient for the task? Do they serve as adequate proxies for the underlying theoretical concepts? Studies of teaching techniques often use grades to measure the amount of learning that is occurring. Are grades a valid instrument for measuring learning?

Relationship validity asks how conclusive the empirical testing was. Was the test appropriate and reasonable? In short, can one conclude based on the empirical test that there is in fact a statistical relationship? Suppose one examines data on energy prices and energy consumption for the decades before and after the Energy Crisis of the 1970s. The price of energy increased dramatically, while energy consumption increased as well. Does that imply that the law of demand does not hold for energy? Casual examination of data on energy prices and consumption lacks relationship validity because it fails to control for other factors that may have affected energy consumption; in other words, the *ceteris paribus* assumption ("other things being equal") does not hold. In the case of the energy crisis, GDP was increasing over the same period, so that actual energy consumption rose despite the higher prices, not because of them.

Causal validity observes that since correlation does not imply causation, can one be sure that the hypothesized causal relationship is valid? Can one be sure that the causation does not go in the opposite direction, or that there is no causation at all, merely a statistical association? For example, consumer spending and personal incomes are highly correlated. Can one conclude that the spending is the result of the income, or the converse? Does yacht ownership cause people to be rich, or does the causation run in the other direction? These questions rarely have easy answers, but careful researchers will ask them nonetheless.

Once you have established the internal validity of a given study, you need to ask if the results can be generalized to other situations, applications, or circumstances. If so, the study has **external validity**. For example, suppose a study found a positive relation between class attendance and GPA for students at small liberal arts colleges. Would the same results hold for courses at a research university taught in large lecture classes and using multiple-choice exams for assessment? Perhaps not. If so, we would conclude that the original study has external validity; if not, we would draw the opposite conclusion.

Note that there tends to be a tradeoff between internal and external validity. As researchers tinker with their model to obtain good statistical results, those results have the tendency to become specific to that case and less generalizable to other situations.

The purpose of empirical testing is to search for evidence in the data that supports your hypothesis. But a good empirical test does more than

that. It also rules out alternative hypotheses. Your research attempts to develop an argument by using logic and evidence in favor of your hypothesis. The problem is that the data in the real world may be consistent with alternative hypotheses as well as with yours. To be successful, you must select an empirical test that adequately discriminates between your hypothesis and alternatives. Statisticians define the **power** of a test as the probability of correctly rejecting a null hypothesis when it is not true. So when you select your empirical test, you should ask yourself the following question: if the test yields the strongest possible statistical results, how confident can I be that my hypothesis is confirmed? In other words, will the test adequately discriminate between your hypothesis and alternatives? If not, then you should consider a more powerful test.³

Consider the following example. In the early 1980s, the Reagan Administration succeeded in reducing personal income taxes by 25 percent. Following the tax cuts, the U.S. economy grew for the remainder of the decade, and this growth was trumpeted as a successful example of supply-side economics. But was it? Two decades earlier, the Kennedy tax cuts and resulting economic expansion were heralded as evidence of the effectiveness of Keynesian economics. One might ask which hypothesis was correct—were the Reagan tax cuts evidence of a supply-side expansion or a Keynesian one?

A key prediction made by the supply-side theory is that tax cuts, by increasing the after-tax return consumers gain by saving, should increase their saving rate. Examining the actual saving rate after the Reagan tax cuts would be a way to test this theory. In fact, however, such an examination is not a very strong test, since the Keynesian theory also predicts an increase in the saving rate (or individual's average propensity to save). The supply-side theory hypothesizes that the higher saving rate stimulates the economy, while the Keynesian theory postulates that the stronger economy promoted by tax cuts raises the saving rate. As a consequence, the empirical test of looking for an increase in the saving rate cannot help us discriminate between the two competing theories.

As you design, conduct, and interpret your empirical testing, you will need to keep these validity issues in mind. We will discuss them again in the next section in the context of specific testing methodologies.

How Does One Analyze Data?

When a researcher begins to think about empirical testing, he or she starts by asking the question: What are the implications (or predictions) of my theoretical analysis? If the hypothesis is true, what evidence should one expect to see? The answer to these questions is called the theoretical pre-

diction of the analysis. Think about how doctors look for symptoms to identify an illness. A woman visits the doctor's office, saying she thinks she has the flu. The doctor thinks to himself: "If she has the flu, she should exhibit the following symptoms: headache, fatigue, fever." So the doctor asks the patient: "What are your symptoms?" It is the symptoms that allow the doctor to make a diagnosis. In the same way, it is the predictions of the theory that allow the researcher to test his or her analysis.

We looked at several examples of this process in Chapter 7. The first example involved a study of the demand for diamond jewelry. The analysis predicted a negative relationship between price and quantity demanded. Another example asked whether parental work hours adversely affect children's cognitive development. If so, work hours and cognitive development should be negatively related.

Once the predictions of the theory are identified, the researcher next asks: Is the evidence of the real world consistent with these predictions? How exactly does the researcher "examine" the evidence to answer this question? We can differentiate between three methods: casual empiricism, simple hypothesis testing, and multiple regression analysis.

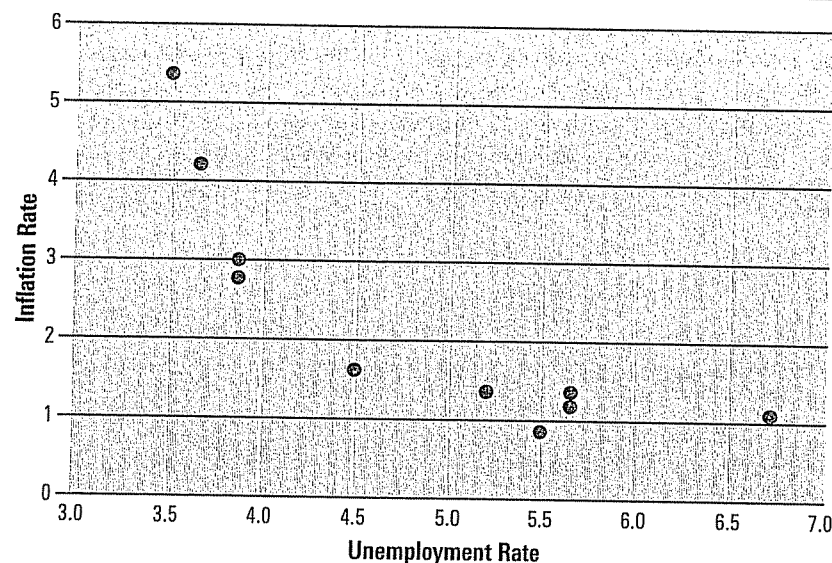
Casual Empiricism

Casual empiricism is the type of analysis one does when constructing a data appendix. It includes printing and graphing data, computing simple descriptive statistics, and visually examining the results. Because it is simple and easy to perform, the researcher can often learn much from this approach, especially when initially looking for hypotheses.

For example, the Phillips curve is the hypothesis that there is an inverse relation between the rates of inflation and unemployment. Figure 10.1 shows the data for these two variables during the 1960s, when the Phillips curve relationship gained prominence. Upon inspection, you can easily see the discernible negative relationship. This important macroeconomic relationship was discovered through casual empiricism.

Note that once you have identified a hypothesis, you should collect a new set of data to test it. You should never test the hypothesis with the same data you used to develop it.

Any sample of data can be described in terms of its **descriptive statistics**. You can think of descriptive statistics as statistics that summarize the underlying data. Descriptive statistics include measures of central tendency and measures of dispersion. A measure of central tendency is the answer to the question: if you had to summarize a sample of data using only one measure, what would you use? The answer is an average. There are three commonly used measures of central tendency or averages. These are the **mean**

Figure 10.1 Phillips Curve

or arithmetic average, the **median** or middle value in the sample, and the **mode** or most common or frequently occurring value in the sample.⁴

How well an average summarizes a data sample depends on the dispersion of the data. There are three commonly used measures of dispersion. The first is the **range** between the highest and lowest values in the data sample. The second is the **standard deviation**, which roughly measures the average amount by which a data point differs from the mean value of the sample. The **variance**, the third commonly used measure of dispersion, is the square of the standard deviation. Any spreadsheet or statistical software package will generate descriptive statistics for a data set.

Most students are familiar with the standard bell-shaped or normal distribution of data. Researchers often assume that their data are normally distributed or approximately normally distributed. If so, all three measures of central tendency will be about the same. If not, the values for the mean, median, and mode of a data sample can differ substantially, so you need to carefully choose which measure of the average to report. Data for personal income is a good example of this. Personal income data tend not to occur in the typical shape of a bell curve. Rather, there are enough high

incomes at the top of the distribution to skew the mean, so the median is a better reflection of the typical income.

In Chapter 9, we noted that computing descriptive statistics for time-series data is not always helpful when those data increase or decrease over time. In these cases, one can look at average growth rates over time, instead of average values. Suppose we return to the Energy Crisis example. Over the time period in question one could ask, what were the percentage changes in energy prices and energy consumption? From 1973 to 1980, for example, the real price of energy increased by 91 percent while energy consumption per dollar of GDP decreased by 23 percent. This implies a negative relationship, as the law of demand suggests.

You can more formally assess the relationship between two variables by examining the **covariance**. The covariance is a measure of how two variables vary together. It is related mathematically to the standard deviation and the variance. Unfortunately, the covariance is difficult to explain intuitively. For that reason, researchers often use a closely related concept, **correlation**. Correlation measures the degree of linear association between two variables. The correlation coefficient, which ranges in value from -1 to $+1$, measures both the direction and strength of the linear relationship. Suppose we graphed energy prices against energy consumption, and drew a line to best fit the data points. The slope of the line would show the (hypothesized) linear relationship between prices and consumption. The slope would be positive or negative. The strength of the relationship would be reflected by the dispersion of the data points—the degree to which the data points are close to or far from the line. If the average dispersion is small, then most of the data points will be found “close” to the line. In this case, we say the correlation between prices and consumption is high. The higher the correlation, the closer the coefficient is to $+1$ for positive relationships or -1 for negative relationships. If the average dispersion is large, then most of the data points will be “far” from the line. In this case, we say the correlation between prices and consumption is low. The lower the correlation, the closer the coefficient is to zero. For example, the correlation coefficient between real energy prices and consumption per dollar of GDP over this time period was -0.95 , which indicates a strong negative linear relationship.

Random Variation in Human Behavior

Data analysis is subject to three major problems. These problems are, first, the effects of random variation in human behavior; second, the fact that relationships between two variables can be concealed by the effects of

other variables; and, third, causal validity, the fact that correlation is not the same as causation. We will explore each of these issues in the following sections.

One of the difficulties faced by research in the social sciences is **random variation** in human behavior. This poses a problem for relationship validity because the effects of random variation can obscure any underlying relationships. Let's illustrate this with an example. Suppose that consumer spending (C) is based *systematically* only on income (Y). Algebraically, this is represented as follows:

$$C = bY, \quad (1)$$

where b is a positive constant showing the relationship between C and Y, that is, the slope of the graph of C against Y.

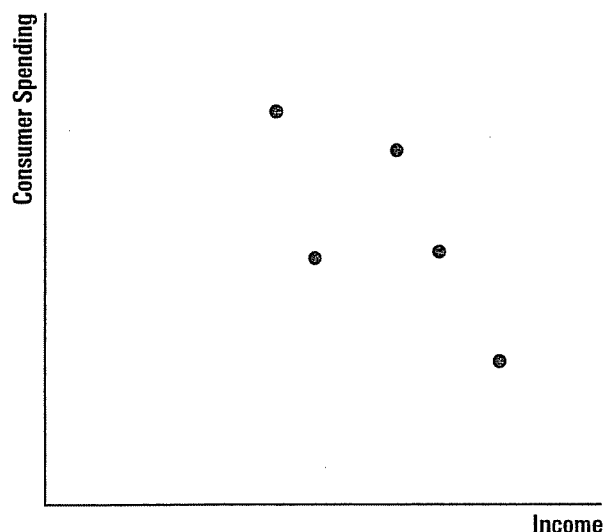
Additionally, however, suppose that in any time period individuals in fact spend a *random amount* more or less than that. As a consequence, consumer spending is actually

$$C = bY + e, \quad (2)$$

where e reflects the random variation in spending.

If we plot the data, they might look something like Figure 10.2.

Figure 10.2 Sample of Data for Consumer Spending and Income



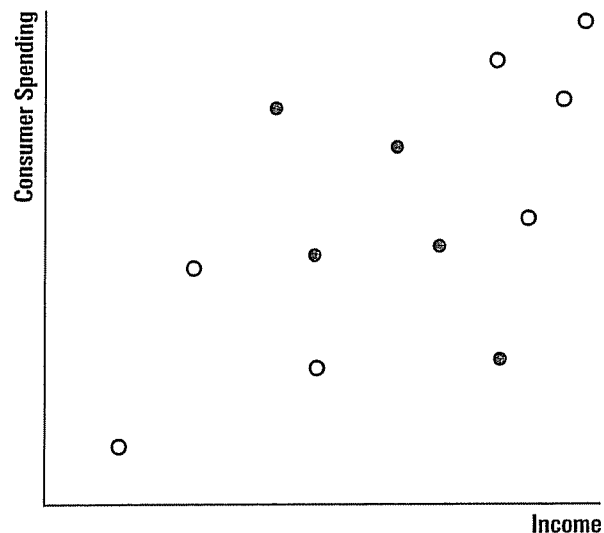
If we imagine a line through the points in Figure 10.2, it would have a negative slope, indicating a negative relationship between C and Y. But are they negatively related? Most social science statistics are based on sample data rather than population data. The key question to ask about any data set is whether or not it fairly represents the underlying population data. Recall that in Chapter 4 we pointed out that for an argument to be persuasive the evidence presented must be representative. This is an example of that point. There are really two issues at stake here: First, is the sample large enough to validly show any underlying relationship in the data? Second, is the sample random? These issues are related, but the second is the more critical.

Suppose consumption and income are positively related, as described by Equation 2. If we take a large random sample of the data, then any data point with a positive error is likely to be matched by another point with a comparable negative error. As a result, the errors will tend to cancel each other out, and the data will show the true underlying relationship. How large a sample is large enough? A common rule of thumb is that a sample size of thirty or more will do the trick. Strictly speaking, the sample should be thirty or more **random** data points. A truly random sample is difficult and expensive to acquire, so typically samples are less than random.⁵ A random sample is also unlikely in macro research, where time-series data are the norm. In either case, bias is introduced into the results.

So where does this leave us? It may be that the data in Figure 10.2 were not a random sample. Thus, the graph may look negatively sloped when the underlying relationship is the opposite. The reason for this anomaly is that some chance exists that *any* data sample will include a nonrandom selection of errors, so the data will appear to show a pattern that does not reflect the underlying population. In other words, the sample of data may not be large enough or random enough for the errors to cancel each other out. As a result, consumer spending will appear to show a negative relationship with income when that appearance is really due to a chance selection of errors. This is called a **sampling error**.

This point is illustrated in Figure 10.3, where the original sample is shown in solid dots, while the remainder of the population is shown in empty dots. Note that a line drawn through the population shows the correct positive relationship between the two variables.

In practice, the problem may be even worse than we've indicated. All data are subject to some degree of measurement error, which is an issue of instrument validity. This measurement error will also be reflected in the e term of Equation 2.

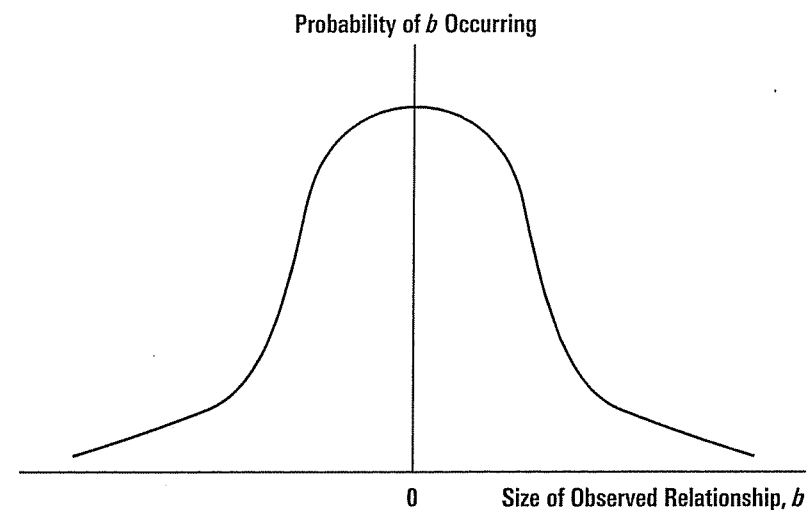
Figure 10.3 Population of Data for Consumer Spending and Income

Statistical Methods

The way to deal with random variation in human behavior is to incorporate statistical methods to determine the likelihood of a sampling error. Let's illustrate how this can be done by continuing the example from the last section, which hypothesizes a relationship between consumption and income (Equation 2).

When employing statistical methods it is critical to differentiate between two concepts: the **null hypothesis** and the alternative or **maintained hypothesis**. The maintained hypothesis is the theoretical prediction of a model. Generally speaking, however, when economists use statistical methods to test a theory, they actually test the null hypothesis. For this reason, the null hypothesis is sometimes referred to as the **statistical hypothesis**, that is, the hypothesis that is actually being tested statistically. If the null hypothesis is rejected by the evidence, then that result supports or is consistent with the maintained hypothesis. In Equation 2, we are estimating the value of b , the relationship between income and consumption. Thus, the null hypothesis, which is actually examined statistically, is that b is zero, while the maintained hypothesis is that b is not zero.

Even if the null hypothesis is true, there is a chance that data for a given sample will show a relationship so that an estimate of b will not be zero. In

Figure 10.4 Distribution of Observed " b " for All Possible Data Samples

fact, the outcomes that we would observe from all possible samples of the data tend to be distributed along a bell curve, as shown in Figure 10.4. You should notice a few things about this bell curve. First, the height of the curve at any point reflects the probability of that observation occurring. Second, the bell curve is symmetric around a value of $b = \text{zero}$, meaning no relationship. Thus, it should be clear that observations closer to zero (i.e., the null hypothesis) will be more likely to occur than observations that are further from zero. This makes sense since we are assuming that the null hypothesis is true. The further from zero that a given observation is, however, the less likely it is to appear. At some point, the probability is so small, the likelihood of observing it is so low, that we reject the null hypothesis in favor of the alternative that the observed relationship is not due to random chance, but rather reflects a real relationship. Such a result would support the maintained hypothesis.

How far away is enough? How certain do you need to be to reject the null hypothesis in favor of the alternative? The answer to this question is called the **level of significance**. The level of significance is the risk that the researcher is willing to take that the null hypothesis will be rejected when it is true. Another way to think about significance level is that it is the probability that the researcher will accept the hypothesis he or she is maintaining when it should be rejected—in other words, that the researcher

NOTES FOR NOVICE RESEARCHERS

Type 1 and Type 2 Errors

Statisticians formalize this discussion of statistical inferences by talking about Type 1 and Type 2 errors. Wrongly rejecting the null hypothesis is called a **Type 1 error**. To illustrate this, consider a person who is on trial for a robbery. In the U.S. legal system, the defendant is assumed to be innocent until proved guilty. Thus, the statisticians would say that the null hypothesis is that the defendant is innocent. Suppose she is indeed innocent, but that the jury convicts her anyway. In this case, they have committed a Type 1 error. Researchers can reduce the likelihood of a Type 1 error by using a smaller level of significance—that is, by requiring a higher standard of proof before accepting that a relationship is valid. Wrongly accepting the null hypothesis is called a **Type 2 error**. To illustrate this case, suppose that the defendant in our example is in fact guilty. If the jury finds her innocent, they would be committing a Type 2 error. Researchers can reduce the likelihood of a Type 2 error by increasing the sample size.

wrongly accepts the maintained hypothesis. Most researchers are willing to use a 5 percent level of significance, though a 1 percent level would be even more convincing. To sum up, if you use a 5 percent level of significance, then you are willing to risk a 5 percent chance that you will reject the null hypothesis when it is true; consequently, there is a 95 percent chance that you will not reject the null hypothesis when it is true.⁶

Simple Statistical Hypothesis Testing

Let's illustrate simple statistical hypothesis testing with an example, using a statistical "*t*-test." Simple statistical hypothesis testing is one case of what is known more generally as **statistical hypothesis testing**. Suppose that we want to know whether the mean value of some sample is significantly different from some standard called the default mean. For example, we could ask if the average unemployment rate among high school dropouts is different from that of the labor force in general. A statistical *t*-test can be used to answer this question.⁷ A *t*-test examines whether the sample mean (in this case, the average unemployment rate among dropouts) is far enough away from the default mean (i.e., the average un-

employment rate among all people) to be considered statistically different. Let's examine how this can be done.

The first step is to compute the estimated *t* statistic using the following formula:

$$t = (\bar{x} - \mu) / (s / \sqrt{n}) \quad (3)$$

where

- \bar{x} is the average unemployment rate among high school dropouts (i.e., the sample mean),
- μ is the unemployment rate in the general labor force (i.e., the default mean),
- s is the standard deviation of the unemployment rate among high school dropouts (i.e., the sample standard deviation), and
- \sqrt{n} is the square root of n , the sample size of the data set.

Note also that s/\sqrt{n} is called the **standard error** of the average unemployment rate among high school dropouts.

The null hypothesis to be examined is that the mean unemployment rate among dropouts is *not* significantly different from that of the labor force in general. This is analogous to our hypothesis of "no relationship" in the previous section. You select the level of significance that you are willing to tolerate (e.g., 5% or 1%). Then you compare the estimated *t* with the critical *t* value for the desired level of statistical significance.⁸ Exact critical *t* values can be obtained from tables in statistics books, such as that in Appendix 10A, or online (e.g., at <http://www.statsoftinc.com/textbook/sttable.html>).

If the estimated *t* exceeds the critical *t*, then you can conclude that the probability of the sample mean being the same as the default mean is low. Thus, it is likely that any observed difference is *not* due to random chance; rather, the observed difference is likely to be real.

Monthly data from January 1993 to October 2003 were obtained from the Bureau of Labor Statistics' webpage (www.bls.gov) to construct this *t*-test. The data were downloaded into an Excel spreadsheet, and the following statistics were computed:

- \bar{x} (the average unemployment rate among high school dropouts) = 8.26%
- μ (the unemployment rate in the general labor force) = 5.28%
- s (the standard deviation of the unemployment rate among high school dropouts) = 1.37
- n (the sample size of the data set) = 130 monthly observations, so $\sqrt{n} = 11.40$
- s/\sqrt{n} (the standard error) = 0.12

Using the formula given in Equation 3, we can construct our observed t statistic:

$$t = (8.26 - 5.28)/0.12 = 24.83$$

The critical t value for this test, using a 5 percent significance level, is 1.98. Since the observed t , 24.83, exceeds the critical t , 1.98, we conclude that high school dropouts have a statistically different unemployment rate than the general population.

Wyrick (1994, 1995) provides a good intuitive explanation of t -tests. He points out that a t -statistic “is a ratio between the observed . . . difference [e.g., between unemployment rates] and the normal variations . . . [in unemployment rates] caused by all factors.” Wyrick argues that “if the ratio is ‘large,’ then the observed . . . difference exceeds the amount that one might attribute to chance fluctuations alone.” Thus, one can reasonably conclude that there is indeed a difference in unemployment rates. Alternatively, “if the ratio is ‘small’ then the observed . . . difference is ‘small’ compared to normal . . . fluctuations, so it would be unwise to conclude that the observed difference is meaningful.”

An Alternative Approach: The P-Value

An alternative approach to performing t -tests is increasingly being used. This is called the p -value. The p -value is the exact or observed level of significance. More formally, assuming that the null hypothesis is true, the p -value is the probability that you would obtain an estimated t score that is as large or larger than the one you obtained, in the direction of the maintained hypothesis. If the observed probability is less than the level of significance, the null hypothesis is rejected.

One advantage of the p -value approach is that it doesn't require consulting a table of critical t -values. Since the p -value approach is equivalent to the t -test, most statistical software packages provide p -values.

Confounding Variables

We stated earlier that the second problem of data analysis is the fact that relationships between two variables can be concealed by the effects of other variables. Statisticians call this problem the confounding of explanatory variables. It affects both casual empiricism and simple t -tests. Even with a random sample of data, it is possible to obtain a sample that looks like Figure 10.2. Suppose that the variable of interest (e.g., consumer spending in the previous example) is affected by something other than the hypothesized cause (i.e., income). Examples of plausible other variables

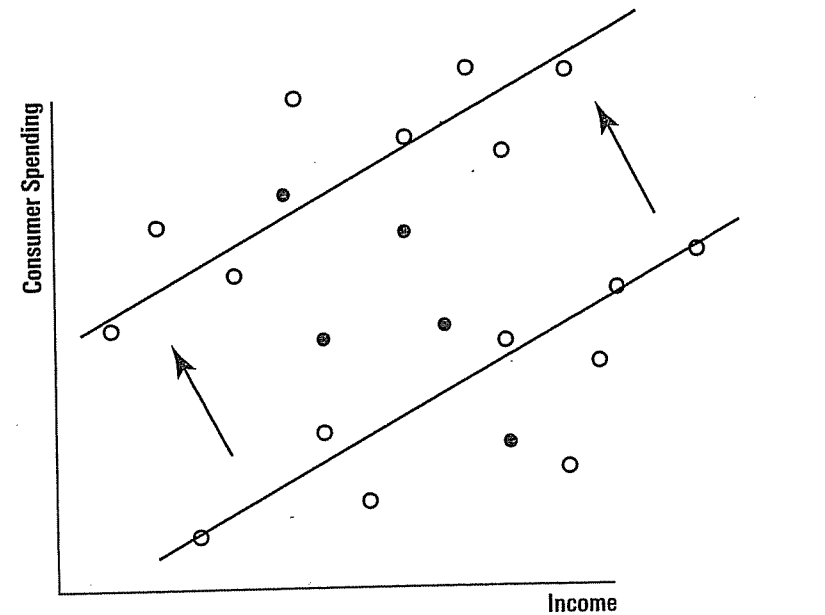
include interest rates or household wealth. Let's amend our equation for consumer spending as follows:

$$C = b_1Y + b_2i + e \quad (4)$$

Now suppose that during the sample period interest rates noticeably decreased, while income didn't. In this case, the data would reflect the effects of interest rates on consumption more than the effects of income. This situation is illustrated in Figure 10.5, where the data sample is reflecting a shift in the consumption function rather than a movement along the function. To accurately measure the relationship between consumer spending and income, we must control for any other variables, like interest rates, that have their own effect on spending. Without this control, visual or statistical examination of the data can be flawed because the impacts of Y and i can't be separated. This is exactly the basis for the *ceteris paribus* (“all other things being equal”) assumption that economists commonly make.

Thus, a control variable is what is held constant in a study. Sometimes, judicious test design can solve this problem. For example, studies

Figure 10.5 The Effects of Interest Rates on the Relation Between C and Y



Using the formula given in Equation 3, we can construct our observed t statistic:

$$t = (8.26 - 5.28)/0.12 = 24.83$$

The critical t value for this test, using a 5 percent significance level, is 1.98. Since the observed t , 24.83, exceeds the critical t , 1.98, we conclude that high school dropouts have a statistically different unemployment rate than the general population.

Wyrick (1994, 195) provides a good intuitive explanation of t -tests. He points out that a t -statistic “is a ratio between the observed . . . difference [e.g., between unemployment rates] and the normal variations . . . [in unemployment rates] caused by all factors.” Wyrick argues that “if the ratio is ‘large,’ then the observed . . . difference exceeds the amount that one might attribute to chance fluctuations alone.” Thus, one can reasonably conclude that there is indeed a difference in unemployment rates. Alternative, “if the ratio is ‘small’ then the observed . . . difference is ‘small’ compared to normal . . . fluctuations, so it would be unwise to conclude that the observed difference is meaningful.”

An Alternative Approach: The P-Value

An alternative approach to performing t -tests is increasingly being used. This is called the p -value. The p -value is the exact or observed level of significance. More formally, assuming that the null hypothesis is true, the p -value is the probability that you would obtain an estimated t score that is as large or larger than the one you obtained, in the direction of the maintained hypothesis. If the observed probability is less than the level of significance, the null hypothesis is rejected.

One advantage of the p -value approach is that it doesn't require consulting a table of critical t -values. Since the p -value approach is equivalent to the t -test, most statistical software packages provide p -values.

Confounding Variables

We stated earlier that the second problem of data analysis is the fact that relationships between two variables can be concealed by the effects of other variables. Statisticians call this problem the confounding of explanatory variables. It affects both casual empiricism and simple t -tests. Even with a random sample of data, it is possible to obtain a sample that looks like Figure 10.2. Suppose that the variable of interest (e.g., consumer spending in the previous example) is affected by something other than the hypothesized cause (i.e., income). Examples of plausible other variables

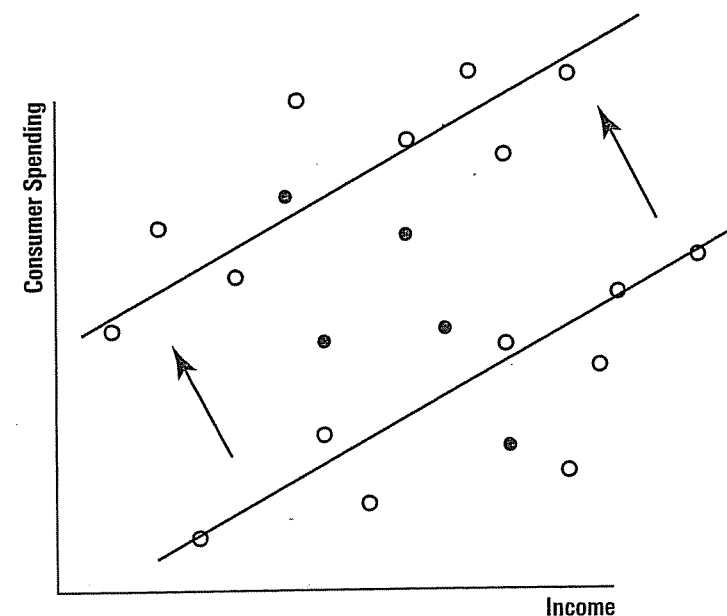
include interest rates or household wealth. Let's amend our equation for consumer spending as follows:

$$C = b_1Y + b_2i + e \quad (4)$$

Now suppose that during the sample period interest rates noticeably decreased, while income didn't. In this case, the data would reflect the effects of interest rates on consumption more than the effects of income. This situation is illustrated in Figure 10.5, where the data sample is reflecting a shift in the consumption function rather than a movement along the function. To accurately measure the relationship between consumer spending and income, we must control for any other variables, like interest rates, that have their own effect on spending. Without this control, visual or statistical examination of the data can be flawed because the impacts of Y and i can't be separated. This is exactly the basis for the *ceteris paribus* (“all other things being equal”) assumption that economists commonly make.

Thus, a control variable is what is held constant in a study. Sometimes, judicious test design can solve this problem. For example, studies

Figure 10.5 The Effects of Interest Rates on the Relation Between C and Y



of energy demand during the 1970s were confounded by the fact that GDP growth caused energy consumption to rise. By looking at energy consumption per dollar of GDP (i.e., Q_d/GDP), it became apparent that, while real energy prices rose and energy consumption also rose, energy consumption per dollar of GDP fell by 20 percent, consistent with the law of demand.

A more general solution is provided by **multiple regression analysis**, which is the most widely used tool of empirical analysis among economists. One way to think about regression analysis is that it is a means for testing hypotheses while controlling for all the other factors that could influence the variable you are interested in. More specifically, regression is a technique for estimating the independent influences of each explanatory variable, and thus statistically maintaining *ceteris paribus*. The details of how to perform and interpret simple multiple regressions will be the focus of Chapter 11.

Causal Validity

The third hurdle of data analysis involves causal validity. Because experimental designs better control for outside factors, they are more powerful for determining causation than are the survey (or nonexperimental) designs that economists typically use. Survey designs usually can only establish association or correlation. The correlation coefficient gives a formal mathematical measure of the relative variation between two variables. One would certainly hope to see a high degree of correlation when assessing a hypothesis. However, the correlation coefficient only indicates the existence of a statistical relationship between the two variables. It doesn't say anything about whether there is a behavioral relationship between the two variables or whether the relationship is merely a statistical accident. Even if there is a behavioral relationship, the correlation coefficient cannot determine whether the causation runs in the hypothesized direction ($C = f[GDP]$), or in the reverse direction ($GDP = f[C]$). In short, it is important to remember that correlation is not necessarily causation. Note especially that regression analysis does not prove causation!

What then does? There is no unambiguous answer here. Existing methodologies for empirical testing, especially using survey research, are not very powerful. In a single research project, probably the most you can hope for is that the empirical evidence will not contradict your hypothesis. In this case, we say that the evidence is consistent with or confirms the hypothesis. Ethridge (1995) points out that the causal implications of survey designs stem from the conceptual analysis rather than from the empirical evidence

NOTES FOR NOVICE RESEARCHERS

What's an Appropriate Empirical Test?

The researcher should choose an appropriate empirical test that is based on the specific objectives of his or her research, considering both the costs and benefits of the testing methodology. This does not mean that the most complex or sophisticated methods are always the best. (Regression may not always be necessary!)

As a general rule, you should use the simplest testing technique that adequately controls for outside factors. Most importantly, remember that the validity of your research design is never higher than the validity of any one part of that design. Thus, if your data are suspect, the most sophisticated statistical techniques cannot improve upon them.

that confirms that analysis. Though survey research cannot prove a hypothesis in the way that experimental research can, multiple survey-based studies using different data sets at different times by different researchers can provide strong evidence of causality in a hypothesis.

SUMMARY

- The purpose of empirical testing is to assess the validity of one's theoretical analysis of an issue or problem.
- The validity of an empirical study has numerous aspects: internal versus external validity, instrument validity, relationship validity, and causal validity.
- A study's validity can be compromised by three hurdles of data analysis:
 1. Random variation in human behavior
Solution: Tests of statistical significance
 2. Dependent variable can be affected by other than the hypothesized independent variable
Solution: Control variables
 3. Correlation is not causation
Solution(s): Sound theoretical development; Multiple studies with different methodologies

NOTES

1. Note that when we say X "causes" Y we are assuming that all other possible explanations for the observed outcome have been controlled for. This is a strong assumption since there is always the possibility that we failed to identify some possible explanation. We will discuss these issues in more detail later in the chapter.
2. One should also consider the reliability and sensitivity of the instruments.
3. Oftentimes, however, lack of adequate data will limit your ability to use a more powerful test. This highlights again the importance of finding a good data set.
4. For even-numbered sample sizes, the median is the number at which half the sample is above and half the sample is below. In other words, it is the number halfway between the two middle numbers in the sample. For odd-numbered sample sizes, the median is exactly the middle number in the sample.
5. Best (2001) points out that social science research often uses "convenience sampling," that is, using the sample that comes to you rather than going to the trouble of developing a random sample.
6. Note that this doesn't say anything directly about when the null hypothesis is false! Statistical testing examines the null hypothesis that there is no statistical relationship. More precisely, it attempts to *disprove* the null hypothesis. The level of significance defines how much confidence one can have that the null hypothesis has been correctly rejected. It does not say anything directly about how much confidence one can place in the maintained hypothesis.
7. For the *t*-test to be appropriate the data should be a normally distributed, random sample.
8. Critical *t* values depend on three factors: the desired level of statistical significance, whether the test is one tailed or two tailed, and the degrees of freedom. A one-tailed test would examine either of two hypotheses: the sample mean (\bar{x}) is less than the default mean (μ), or the sample mean is greater than the default mean. By contrast, a two-tailed test would examine the hypothesis that the sample mean (\bar{x}) is not equal to the default mean (μ). Thus, it is a more general test. The degree of freedom is defined as $n - 1$, where n is the sample size. In our example, the degrees of freedom are $130 - 1 = 129$.

EXERCISES

1. Collect twenty years of annual data for U.S. inflation rates and the growth rate of the money supply (M1). For the inflation rate and the money-growth rate compute the mean, median, and mode, as well as the standard deviation and range of values. Plot the two variables on a graph. Do you see any relationship between the two variables? Compute the correlation between inflation and money growth.
2. What are the key variables in your research hypothesis, that is, your dependent variable and the hypothesized variables that affect it? Collect twenty data points for each variable. Compute the mean, median, and mode, as well as the standard deviation and range of values for each variable. Plot the dependent variable with each key independent variable on a graph. Do you see any relationship between the two variables? Compute the correlation between the two variables.
3. Is the unemployment rate in your state statistically different from the national unemployment rate over the past ten years? Construct a *t*-test to assess this hypothesis.
4. Think about your research hypothesis. What variables other than the key variables that you identified in exercise 2 might influence the dependent variable and thus need to be controlled for in your empirical testing?

SUGGESTIONS FOR FURTHER READING

- Best (2001)*—Very thoughtful monograph on how data series are constructed and the ways they can distort what they purport to measure.
- Huff (1954)*, especially Chapters 3, 4, and 8—Excellent intuitive explanation of statistical significance.
- McCuen, Johnson, and Davis. (1993)*—Very readable introduction to technical writing for scientists; see especially Chapter 4, Communicating Technical Data and Statistics.
- Phillips (1996)*—Very thoughtful, intuitive guide for users of statistical methods.
- Wyrick (1994)*—Guide to regression analysis for undergraduates.

Table of Critical t -Statistics

		Level of Significance		
		One-Sided Test		
Degrees of Freedom		.10	.05	.01
1		3.078	6.315	31.82
2		1.886	2.920	6.965
3		1.638	2.353	4.541
4		1.533	2.132	3.747
5		1.476	2.015	3.365
6		1.440	1.943	3.143
7		1.415	1.895	2.998
8		1.397	1.860	2.896
9		1.383	1.833	2.821
10		1.372	1.812	2.764
11		1.363	1.796	2.718
12		1.356	1.782	2.681
13		1.350	1.771	2.650
14		1.345	1.761	2.624
15		1.341	1.753	2.602
16		1.337	1.746	2.583
17		1.333	1.740	2.567
18		1.330	1.734	2.552
19		1.328	1.729	2.539
20		1.325	1.725	2.528
21		1.323	1.721	2.518
22		1.321	1.717	2.508
23		1.319	1.714	2.500
24		1.318	1.711	2.492
25		1.316	1.708	2.485
26		1.315	1.706	2.479
27		1.314	1.703	2.473
28		1.313	1.701	2.467
29		1.311	1.699	2.462
30		1.310	1.697	2.457
∞		1.282	1.645	2.326

Source: This table was generated using the Quantile, StudentTDistribution, and NormalDistribution commands in *Mathematica*.

Introduction to Regression Analysis

"You can use all the quantitative data you can get, but you still have to distrust it and use your own intelligence and judgment."

ALVIN TOFFLER

In the last chapter, we noted that regression analysis is the most widely used method by which economists test hypotheses. For that reason, it deserves a chapter of its own. This chapter won't substitute for an econometrics course, however. It is more like a course in music appreciation; it won't make you a musician, but it will enable you to understand and respect certain types of music and perhaps even make some music yourself. The chapter has two modest goals:

- To enable you to understand published regression results well enough to determine whether they support a hypothesis, and
- To provide you with a guide to performing simple regressions to test your own research hypotheses.

Steps in Regression Analysis

Regression analysis is a means of testing hypotheses while controlling for other factors that could influence the variable you are primarily interested in. A regression analysis can be broken down into five steps, as illustrated in Figure 11.1. We explore these steps in more detail in the sections that follow.

Figure 11.1 Steps in a Regression Analysis

Step 1: State the Hypothesis

Step 2: Test the Hypothesis, i.e., Estimate the Relationship

Step 3: Interpret the Test Results

- To what extent do the coefficient estimates conform to the maintained hypothesis identified in Step 1?
- Are the coefficient estimates statistically significant?
- Are they economically significant?
- Are the coefficient estimates plausible for the real world, consistent with economic theory, and within the range of previous estimates?
- How "good a fit" is the overall regression model?

Step 4: Check for and Correct Common Problems of Regression Analysis

Step 5: Evaluate the Test Results

- On balance, to what extent do the results support the hypothesis?

Step 1: State the Hypothesis

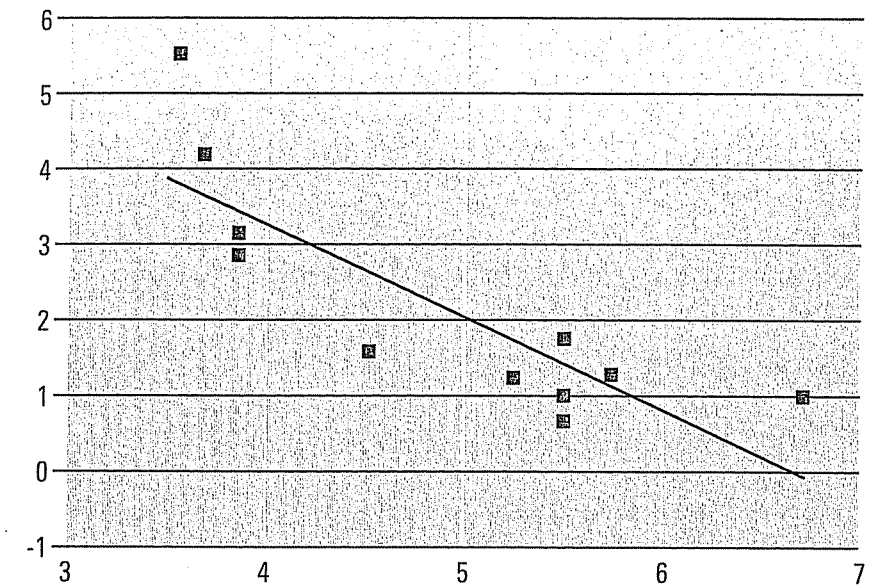
The first step in a regression analysis is to state the maintained hypothesis. Suppose you observed the substantial increase in equity prices during the 1990s and wondered to what extent that increase might have contributed to the consumer spending that powered the economic expansion in the same period. You could hypothesize that stock market wealth (SMW) has a positive effect on consumer spending:

$$C = f[\text{SMW}], \text{ where } f' > 0 \quad (1)$$

Researchers call C the **dependent variable**, in other words, the concept we are trying to explain. By contrast, SMW is an **explanatory or independent variable**, that is, a concept we hypothesize that explains or causes the dependent variable.

Traditionally, analysts have assumed that a linear relationship exists between the dependent and independent (explanatory) variables.¹ This assumption is justified on two bases: first, nonlinear estimation is extremely difficult and time consuming to perform, and, second, even if a relationship is nonlinear, one can often approximate it with a linear relationship. The latter point is illustrated in Figure 11.2, which shows a plot of inflation versus unemployment. The function that best fits the data would be a curve rather than a line, but the line comes fairly close to representing the

Figure 11.2 Linear Approximation to a Nonlinear Function



function. However, the growing availability of computer programs that can perform nonlinear calculations has made nonlinear regressions more common. Nonetheless, linear regression is still the norm, and we will use it here.²

If we assume a linear relationship between stock market wealth and consumer spending, we can express the hypothesis in algebraic form as follows:

$$C = a + b \text{ SMW} \quad (2)$$

If you think about this in terms of a line, then b shows how much C changes as a result of a change in SMW, assuming nothing else changes at the same time. In other words, b shows the slope of the line. Similarly, a represents the vertical intercept of the line. Mathematically, it shows what C will be when SMW is zero; economically, it shows the effects of everything besides SMW on C . Because there are only two explicit variables in this hypothesized relationship, we call it a bivariate regression.

So far, we have not discussed one of the major advantages of regression analysis: it gives us the ability to control for the effects of confounding variables. Though we are primarily interested in the relationship between stock market wealth and consumer spending, economic theory suggests

other variables may affect consumer spending too, for example, income (Y) and perhaps other forms of wealth (OW). We can easily incorporate additional variables into the model as follows:

$$C = a + b_1 \text{SMW} + b_2 Y + b_3 \text{OW} \quad (3)$$

Most regression studies, like this one, are **multiple** (or multivariate) **regressions**, since the dependent variable is influenced by multiple explanatory variables.

Once we have formally stated the hypothesis in the form of Equation 3, it is important to identify our theoretical predictions. Note that there is a maintained hypothesis (or theoretical prediction) for each explanatory variable in the model. In our stock market wealth example, the main hypothesis predicts that higher stock market wealth should lead to higher consumer spending. Thus, b_1 should be a positive number. The researcher should also identify the predictions for the control variables. For example, one might expect that the effects of income and other forms of wealth on consumer spending, shown by the values of b_2 and b_3 , would be positive. It is good scientific practice *never* to run a regression without explicitly identifying the results that would confirm the hypothesis.

Step 2: Test the Hypothesis; i.e., Estimate the Relationship

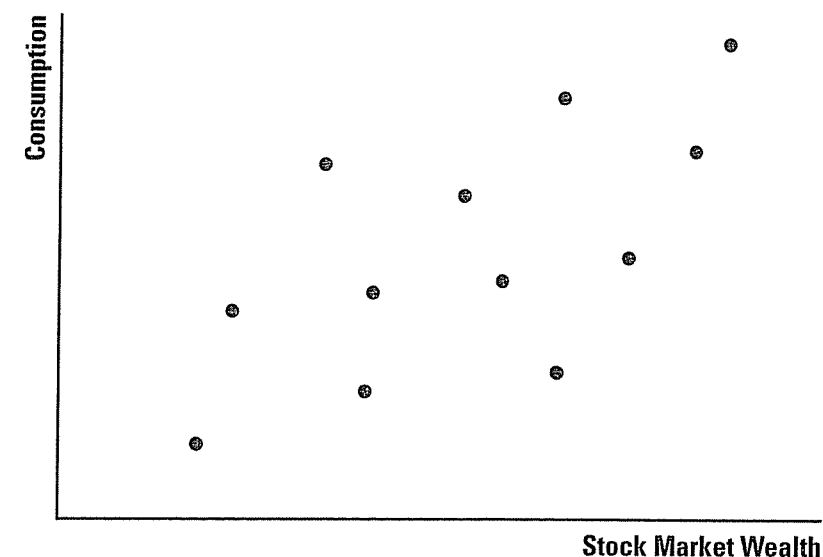
The next step in regression analysis is to analyze the data on the dependent and explanatory variables for any possible relationship between them. We observed in Chapter 10 that even if the maintained hypothesis is true, each observed data point includes an error caused by random chance, measurement error, or human subjectivity. Thus, the multiple regression model becomes the following:

$$C = a + b_1 \text{SMW} + b_2 Y + b_3 \text{OW} + e. \quad (4)$$

A plot of the data for C and SMW, shown in Figure 11.3, displays a scatter of data points with an upward slope.³

Though Figure 11.3 seems to indicate a positive relationship, this conclusion is complicated by two factors. First, we don't know what the errors are for each observation. Second, we only have a sample of the data, rather than all the data, so it's unlikely that the errors will cancel out. As a consequence, we can't simply measure the underlying relationship. Instead, it is necessary to *estimate* the relationship. This is what regression is all about. More specifically, we make certain assumptions about the error, which allow us to make estimates of the underlying relationship.⁴ Violating any of the standard assumptions will require us to use different estimation techniques.

Figure 11.3 Plot of Data for Consumption and Stock Market Wealth



Regression analysis asks the question: what line “best” fits the data (for example, in Figure 11.3). Recall that the equation of any line is defined by specific values for the slope and intercept. Thus, the regression process chooses values for the slope and intercept so that the estimated relationship best fits the data. The various estimation techniques used by economists deal with defining “best” given the specific data at hand.

In short, the purpose of regression analysis is to generate estimates of the relationship between the dependent variable (e.g., C) and each of the explanatory variables (e.g., SMW, Y, OW). Because each of these relationships is estimated with all the others held constant, they are analogous to partial derivatives in calculus. These estimates are called **estimated parameters** or **estimated coefficients**, and they are the primary product of the regression analysis. Fortunately, we no longer have to do this by hand. Instead, a variety of computer programs are available (e.g., Excel, E-Views, SPSS, SAS) that perform regression analysis.

Let's look at an example. Figure 11.4 shows the results gained from estimating the relationship between real consumer spending (CONS) and two independent variables: stock market wealth as measured by the Wilshire 5000 stock market index and real personal disposable income (REAL_PDI). We will use this as our example as we discuss various aspects of regression analysis.

Figure 11.4 Sample Regression Results

Dependent Variable: CONS				
Sample: 1990:4 2002:1				
Observations: 46				
Variable	Coefficient	Std. Error	t Statistic	Prob.
C	-781.48	178.73	-4.37	0.0001
W5000	0.0096	0.0067	1.43	0.1587
REAL_DPI	1.0419	0.0389	26.8	0.0000
R-Squared: 0.9939				
Adjusted R-Squared: 0.9936				
Durbin-Watson: 1.4391				
F: 3519.2				

Whatever statistical software you use, the results will include the coefficient estimates, shown in the second column of Figure 11.4, and a variety of test statistics, which can be used to evaluate the text results. We will discuss how to do this in the next section.

Step 3: Interpret the Test Results

Probably the most important step in regression analysis, and the step novice researchers put the least effort into, is interpreting the results. Interpreting regression results is *not* a mere mechanical exercise. To be valid, regression requires judgment on the researcher's part. Recall that the purpose of regression is to assess the validity of your theoretical analysis. Is the empirical evidence consistent with your maintained hypotheses? Does it support the hypotheses? To answer these questions, it is necessary to evaluate the regression results. As you do so, it is important to remember the point made in Chapter 10—researchers typically test the null hypothesis, that there is no relationship. Thus, the statistical hypothesis is the opposite of the theoretical prediction.

Start with Your Theoretical Predictions To begin interpreting the results, you need to remind yourself what evidence you hoped to see to confirm your hypothesis. We introduced this in Step 1. Did you expect to see a positive or negative relationship between each explanatory variable and the dependent variable? Did you have any expectations about the magnitudes of each coefficient? The simple quantity theory of money illustrates this well, since it makes predictions not only about whether the coefficient estimates are positive or negative but also about their size.

Figure 11.5 Sample Regression Results

	a	b
Theoretical Prediction	0.0	1.0
Estimated Value	0.004	0.91
Estimated t Statistics	0.2	2.4
Estimated p Value	0.84	0.02

The quantity theory was the theory of aggregate demand used by the classical economists. The classical economists assumed that aggregate demand could have no effect on the levels of employment or output in the economy. As such, they hypothesized that the inflation rate (π) should be same as the growth rate of the money supply ($\% \Delta M_s$) and that, assuming no economic growth or changes in the velocity of money, the rate of inflation would be affected by nothing else. This hypothesis can be expressed in the form of an equation:

$$\pi = a + b \% \Delta M_s \quad (5)$$

In short, the quantity theory predicted that a (the vertical intercept in a regression) would be zero and that b would be positive one.

To What Extent Do the Coefficient Estimates Conform to Your Theoretical Predictions?

Once you have clearly stated your theoretical predictions, it is time to examine the coefficient estimates. Consider the predicted and estimated values of the coefficients for the quantity theory of money, as shown in Figure 11.5.

What are the algebraic **signs** of the estimated coefficients? In Figure 11.5, we see that the coefficient on the money growth term (i.e., b) is positive (+0.91). But we cannot take that positive sign as conclusive evidence. Recall our discussion from Chapter 10 that the empirical results from a given data sample could be the result of random chance rather than of the hypothesized relationship.

Regression software typically reports several test statistics to help the user assess this possibility. For each coefficient estimate b , the software computes a t statistic using the formula we introduced in Chapter 10:

$$t = (\bar{x} - \mu) / (s / \sqrt{n}), \quad (6)$$

where b takes the role of x in this equation, and μ is now the true value of the coefficient. The specific hypothesis tested by the regression software is

NOTES FOR NOVICE RESEARCHERS

Quick Rule of Thumb for the Critical t

A commonly used rule of thumb for the critical t statistic is that it is 2.0. This is because as the sample size increases the critical t declines from a little above 2.0. You should confirm this by looking at the t tables in Appendix 10A. Thus, for almost any level of significance and sample size, if the estimated t is greater than 2.0 then it will be greater than the exact critical t , and therefore the estimated coefficient will be statistically different from zero.

whether the estimated coefficient is statistically different from zero. Thus, using the language from Chapter 10, the null hypothesis is $b = 0$, while the theoretical prediction (or the maintained hypothesis) is $b \neq 0$. Substituting the null hypothesis into the preceding equation yields:⁵

$$t = (b - 0) / (s / \sqrt{n}) \quad (7)$$

$$t = b / SE, \quad (8)$$

where:

- b is the estimated coefficient between consumer spending and income, and
- SE is the standard error of that estimated relationship, defined as the standard deviation of the independent variable divided by the square root of the sample size of the data set.

If the estimated t exceeds the critical t for the desired significance level, then we reject the null hypothesis and conclude that the coefficient is statistically different from zero (or **statistically significant**). Thus, the reported algebraic sign is likely valid.

By contrast, if the estimated t does not exceed the critical t , we can't reject the null hypothesis. Thus, it is likely that any observed relationship is simply due to random chance, so the reported algebraic sign cannot be accepted as valid.

Regression software also computes a p value, which provides an alternative approach to evaluating t statistics. The p value is the probability that you would obtain the estimated t score that you obtained if the null hypothesis is assumed to be true. If the p value is sufficiently small, or more precisely smaller than the level of significance, one can reject the

NOTES FOR NOVICE RESEARCHERS

Assessing the Statistical Significance of an Estimated Coefficient

Our hypothesis in the example shown in Figure 11.4 was that consumer spending during the 1990s was driven by the booming stock market. To assess this hypothesis, let's examine the coefficient estimate for "W5000," which measures the relationship between the Wilshire 5000 stock market index and consumer spending. The estimated coefficient on W5000 is 0.0096, and the standard error is 0.0067. Using the formula given in Equation 8, we can confirm that $t = b / SE = 0.0096 / 0.0067 = 1.43$. The degrees of freedom for this test are $n - 1 = 46 - 1 = 45$. For a 95 percent level of confidence and a one-tailed test, the critical t is 1.68.⁶ Since the estimated t is less than the critical t ($1.46 < 1.68$), we conclude that the estimated coefficient is *not* statistically different from zero. Thus, our hypothesis is not confirmed by the data—that is, the hypothesis fails.

null hypothesis and conclude that the coefficient estimate is statistically significant.

For the example in Figure 11.5, suppose the critical t is 2.0. Since the estimated t is 2.4, which is greater than the critical t , we can conclude that a positive relationship does appear to exist between inflation and money growth, which is consistent with the hypothesis of classical economic theory. Alternatively, using the p value approach, since the p value is 0.02, which is less than the level of significance, 0.05, we reject the null hypothesis and conclude that the estimated coefficient is statistically different from zero.

Now what about the **size** of the coefficient estimates? Some hypotheses make very specific predictions about the magnitudes of the coefficients. For example, the quantity theory predicts that the constant term, a , will be zero and that b will be one. Even if a hypothesis does not make specific predictions about the size of the relationships, the coefficient estimates are important for several reasons, which we will discuss shortly.

In assessing the magnitudes of coefficients, we must consider several issues. In empirical work, it is very unlikely that the estimated coefficients will *exactly match* the predictions. Indeed, if they did, you should be very

suspicious! After all, we are examining the behavior of human subjects. Look back at Figure 11.5. The estimated coefficients are roughly the same size as the hypothesized ones. The question is, how close is close enough? One way to judge this is by constructing a test to see if the estimated coefficient is statistically different from the theoretical prediction. To do this, start with the formula for the t statistic given in Equation (6), which we rewrite here:

$$t = (\bar{x} - \mu) / (s / \sqrt{n}) \quad (6)$$

The null hypothesis is that the two coefficients are equal; thus, substituting the null hypothesis into the formula yields the following:

$$t = (\hat{b} - b) / (s / \sqrt{n}) \quad (9)$$

or

$$t = (\hat{b} - b) / SE \quad (10)$$

where:

- \hat{b} is the estimated coefficient,
- b is the predicted coefficient,
- SE is the standard error of the estimated coefficient.

In our example, the test statistic is $(0.91 - 1.0) / 2.6 = -1.6$.⁷ Since this estimated t is less than the critical t of 2.0, we can conclude that the estimated coefficient is not statistically different from the theoretical coefficient. In other words, the estimated coefficient is close enough. As such, this result, like the test of the algebraic sign earlier, supports the hypothesis.

Note that the same test can be conducted on the constant term. In this example, since the predicted coefficient is zero, you can use the default test reported by the regression software without making any calculations.

Statistical versus Economic Significance

Performing a statistical analysis of the estimation results is the beginning rather than the end of the assessment. Even when the theory makes no specific prediction about the size of the coefficient, to be useful in the real world the analyst needs to know the magnitude of the relationship. Consider the following equation:

$$GDP = a + b INT \quad (11)$$

where INT is the interest rate. In the context of monetary policy, it matters a great deal whether b is -0.001 or -1000 .

McCloskey (1998) argues that economists often confuse statistical significance with scientific or **economic significance**. Suppose that in our ear-

lier regression, the estimated t statistic for b was greater than the critical t , so one could conclude that the coefficient is statistically different from zero. But if the coefficient is numerically small (e.g., -0.001), then the explanatory variable (in this case, the interest rate) may not matter much in practical terms. In other words, even though it is statistically significant, it may not be a very important determinant of the dependent variable, GDP.

This point is reinforced by Kennedy (2003) who observes that since nearly any relevant variable will have some (i.e., small) measured effect on the dependent variable, any variable can be made statistically significant if the sample size is large enough. But what truly matters is the size of the effect.

McCloskey also shows the converse argument: that a variable can fail the statistical test but still be economically important. The theory of purchasing power parity (PPP) asserts that the exchange rate between two currencies should be equal to the ratio of the two nations' aggregate price levels:

$$ER = P_{us} / P_{japan} \quad (12)$$

where:

- ER is the exchange rate (U.S. dollars/Japanese yen).
- P_{us} is the consumer price index in the United States.
- P_{japan} is the consumer price index in Japan.

Suppose a test of the theory of PPP yields a coefficient of 0.9, which is shown to be statistically different from the theoretical hypothesis of 1.0 (perhaps because of a large sample). Should one conclude that the theory fails? Careful consideration of the theory could lead to the conclusion that real-world imperfections such as transportation costs, which the theory ignores, prevented the coefficient from being closer to the theoretical one. In short, even though the estimated coefficient is statistically different from the null hypothesis, it may be economically close enough, meaning that the behavior posited by the theory is in fact occurring.

Are the Coefficient Estimates Plausible?

When you examine a coefficient estimate, you should ask several questions. Suppose that you have estimated the demand for new cars as a function of the price of new cars, as well as a number of other explanatory variables:

$$Q_d = a + b_1 P + b_2 X \quad (13)$$

where:

- Q_d is the quantity of new cars demanded,
- P is the price of new cars, and
- X is the other explanatory variables.

NOTES FOR NOVICE RESEARCHERS

The Largest Estimated Coefficient Is Not Always the Most Important!

Researchers sometimes think that the variable that has the largest estimated coefficient is the most important. After all, it has the largest impact on the dependent variable. Or does it? Suppose that one variable is defined in millions while another is defined in billions. If the coefficient on the latter is smaller than the coefficient on the former, it may still have a larger impact on the dependent variable than the former if it is less than one thousand times smaller. Alternatively, suppose that one variable is measured in billions of dollars but the other is measured in percentage points (e.g., interest rates). It's not obvious, then, that the larger coefficient has the larger effect, since a one-unit change in billions of dollars is not the same as a one-point change in the interest rate. Another common error is to think that the variable that has the largest t -statistic is the most important variable. As we noted earlier, a small coefficient with a large t statistic has only a small effect, so don't make this mistake.

What are the units of the parameter estimates? If you don't know what the units are it is difficult to interpret the meaning of the parameter estimates. Suppose price is measured in thousands of dollars, and the quantity demanded is measured in thousands of cars. Then, from Equation 13, b_1 is $\Delta Q_d / \Delta P$ = thousands of cars / thousands of dollars = cars per dollar. You should always use meaningful units to report regression results!

Given the units, are the parameter estimates plausible? Are they consistent with common sense? Suppose b_1 is estimated to be $-5,226,371,462$. This implies that a one-dollar decrease in price causes an increase in new car purchases of roughly 5.2 billion cars. Given that total U.S. car production in 1999 was approximately 5 billion cars, this coefficient estimate is suspect. It is not reasonable to think that demand for cars would double if the price fell by merely one dollar.

Getting implausible parameter estimates suggests an econometric problem of one type or another. For example, the model might omit a key variable or the underlying relationship could be nonlinear. We will turn to these issues shortly, in Step 4.

NOTES FOR NOVICE RESEARCHERS

Assessing the Plausibility of an Estimated Coefficient

The example in Figure 11.4 included income as a determinant of consumer spending. Therefore, the estimated coefficient on income is the marginal propensity to consume. The coefficient estimate was 1.04, which, strictly speaking, is outside the range of theory. But is it implausible? There is a great deal of evidence to suggest that consumers went on a spending spree during the 1990s. The estimate of 1.04 is not significantly different from the upper limit of the theoretical range, so one could conclude that it is plausible. By contrast, if the estimated coefficient were five or ten times larger or smaller, it would not be plausible.

In more general terms, one way to evaluate the plausibility of parameter estimates is to convert them into elasticities.⁸ For a linear demand equation, the estimated coefficients are $\Delta Q_d / \Delta P$, $\Delta Q_d / \Delta Y$, and so on. The price elasticity of demand is defined as $\% \Delta Q_d / \% \Delta P$, or $\Delta Q_d / \Delta P * P / Q_d$. To obtain an average value for the elasticity, you can multiply the estimated coefficient by the ratio of the mean price in the data sample to the mean quantity demanded. Other demand or supply elasticities can be computed in a similar way. For example, to estimate the average income elasticity of demand, you can multiply the estimated coefficient on the income term by the ratio of the mean income in the data to the mean quantity demanded.

Another question to ask when assessing the values of estimated coefficients is whether or not they are consistent with economic theory. In other words, does theory imply anything about the size of a coefficient? For example, in estimating a consumption function, the parameter on income is the marginal propensity to consume, which theory says should be a positive fraction (i.e., $0 < b < 1$). Is your estimated coefficient a positive fraction?

Additionally, you should ask how the parameter estimates compare with previous estimates by other scholars. Are they near that range? If they are too far off, that might indicate an econometric problem, which should be investigated. If your coefficient estimates "pass" all of these tests, you can treat them as valid.

You should note that it is not uncommon to have somewhat mixed results in which some parameter estimates are consistent with your a priori expectations (i.e., close enough), while others are not. These are imperfect tests, after all, and the data are rarely ideal. In these cases, the analyst needs to exercise judgment. Do the explanatory variables involving the *specific* hypothesis support the theory, while control variables may not? Are most of the variables consistent with the theory?

How “Good a Fit” Is the Regression Model?

We noted earlier that regression analysis selects parameter estimates in order to create a regression line that “best” fits the data. Most regression software computes test statistics to help the user determine how well this goal is achieved. The two most common of these are the R^2 (or adjusted R^2) and the F statistic.

The R^2 statistic is an estimate of the proportion of the variation in the dependent variable that is explained by the model. The higher the proportion, the better the model explains the variation in the dependent variable. Note that this is not identical with “proving” your hypothesis. We will explain this point in more detail shortly. Because of the way that R^2 is computed, whenever you add more explanatory variables, the R^2 will increase. In fact, the R^2 will increase even if the variable you add does not significantly affect the dependent variable. But this is cheating! The adjusted R^2 , usually called \bar{R}^2 corrects for this. We won’t distinguish between the two from here on, but if you have a choice you should rely on the adjusted R^2 .

The F statistic tests the hypothesis that *all* estimated coefficients are jointly (or collectively) equal to zero; that is, it tests the hypothesis that $b_1 = b_2 = \dots = b_n = 0$, for all n coefficients. This works much like the t test described previously. If the estimated F statistic is greater than the critical F obtained from statistical tables, then we can reject the null hypothesis and conclude that all coefficients are statistically significant. Thus, the model as a whole has some validity. Note that the F statistic examines the same question as the R^2 does: namely, how well the model as a whole works at explaining the data.

Statistical significance is a good thing, but it is often overrated. For example, it is common for novice researchers to believe that the single most important result of a regression analysis is a “high” R^2 . This is, frankly, not true. In the context of testing hypotheses, R^2 can provide useful diagnostic information, but only indirectly. (By contrast, in the context of forecasting, R^2 is more directly relevant.) A low R^2 may be an indication that one or more explanatory variables are missing from the model; thus, in-

cluding them could improve the accuracy of the results. Kennedy (2003) points out that a high R^2 may be caused by the peculiarities of a specific data sample—though it fits the current sample well, it may fit other samples poorly. Thus, internal validity may be achieved at the cost of external validity. In short, statistical significance is neither necessary nor sufficient to confirm one’s hypothesis. Consider another example. If the estimated coefficient has the wrong sign but is statistically significant, the hypothesis fails. This issue will be discussed more in Step 5, later in this chapter.

One final note: R^2 for cross-section data is generally less than R^2 for time-series data. Econometricians typically consider a time-series regression to be “good” if it results in an R^2 of 0.8 or higher. By contrast, a cross-section regression is considered “good” if it has an R^2 of only half that: 0.4 or above.

Step 4: Check For and, If Necessary, Correct Common Problems of Regression Analysis

We began this chapter by framing the question that is the focus of the chapter: how does one use regression analysis to empirically test a hypothesis? The answer is that regression provides estimates of coefficients that show the signs and magnitudes of a hypothesized relationship, when controlling for other relevant factors.

However, the validity of the ordinary-least-squares (OLS) regression methodology, which we discussed earlier in this chapter, depends on a number of technical assumptions holding true. In practice, it is not uncommon for one or more of these assumptions to be false. When that happens, the resulting parameter estimates are not ideal, which means that the process may not adequately test the hypothesis.⁹ So, it is always prudent to investigate this possibility. Indeed, you should do so before you invest a lot of time and effort in interpreting the initial regression results you obtained. Typically, econometric software packages provide summary statistics to enable the researcher to do this.

Let’s consider five common problems of regression analysis: autocorrelation, heteroskedasticity, specification error, multicollinearity, and simultaneous equations bias. We will define each problem, suggest how to diagnose it, and offer possible treatments. Note that this chapter, as a brief introduction to regression, will not discuss regression problems in great depth. For a more complete treatment, see any econometrics text, such as those listed in Suggestions for Further Reading at the end of the chapter.

Problem 1: Autocorrelation The ordinary-least-squares regression model is based on the premise that the underlying relationship between the

dependent and explanatory variables is linear (or additive). In other words, the regression equation can be expressed as follows:

$$Y = b_0 + b_1X_1 + b_2X_2 + \dots + b_nX_n + e \quad (14)$$

where Y is the dependent variable, the X 's are one or more explanatory variables, and e is the error term. The model further assumes that the error in each observation of the data sample is independent of the error in all the others. For example, what this means is that if one error is positive the next error is no more likely to be positive than negative.

Suppose this were not the case. If a positive error was more likely to be followed by another positive error, or a negative error was more likely to be followed by another negative error, then we say that the errors are auto (or serially) correlated. **Autocorrelation** means that the errors are dependent on or correlated with each other. (Serial correlation refers to autocorrelation in time-series data; it means that the error in one time period is dependent on or correlated with the error in the next period.)

The consequence of autocorrelation is that the standard error measured by the regression software will be biased downward, so that the reported t scores will be artificially inflated. This means that tests of significance may indicate that the explanatory variables are statistically different from zero when they are not.

Autocorrelation suggests that something systematic in the determination of the dependent variable has been omitted from the regression equation. One possibility is that a relevant explanatory variable has been left out. This would be an example of a **specification error**, that is, using an incorrect model. In this case, the solution would be to add the missing variable to the estimation.

True autocorrelation, however, occurs when the problem resides in the errors themselves rather than being a symptom of something else, for example, some sort of inertia that allows a disturbance to spread either across time or space. One example of this might be a lagged adjustment of investment to a change in interest rates. Another could be when a downturn in the economy spills over from one state to a neighboring one. Whatever the cause, first-order autocorrelation is modeled as follows:

$$e_t = \rho e_{t-1} + u_t \quad (15)$$

where ρ is a fraction that indicates the extent to which the error in one observation spills over to affect the error in the next observation, and u_t is a random component of the error.

The easiest way to check for first-order autocorrelation is to use the Durbin-Watson statistic (DW), which is computed by most econometric

NOTES FOR NOVICE RESEARCHERS

Quick Rule of Thumb for the Durbin-Watson

In an earlier note, we observed that a quick rule of thumb is to use 2.0 as the critical t in hypothesis testing. Here is a rule of thumb for assessing the possibility of autocorrelation: The closer the Durbin-Watson is to 2.0, the less likely it is that autocorrelation is a problem.

software packages by estimating ρ in Equation 15. When $\rho = 0$ (meaning no first-order autocorrelation), $DW = 2.0$. This is the null hypothesis. The greater the difference between DW and 2.0, the more likely it is that autocorrelation exists. Specifically, there are two critical DW statistics: DW_U and DW_L (for upper and lower, respectively). If the estimated DW is greater than DW_U , then one can conclude that DW is not statistically different from 2.0, and thus, one cannot reject the null hypothesis of no autocorrelation. This is the desired result. If DW is less than DW_L , the null hypothesis is rejected in favor of the alternative of first-order autocorrelation. If DW falls between DW_L and DW_U , it falls in the ambiguous range and requires the researcher to make a judgment about how to proceed.

Most econometric software gives users the ability to correct for first-order autocorrelation by estimating the relationship and adding it to the regression model. As a result, when the model is estimated the t scores will be correctly reported and the coefficient estimates correctly assessed. Consult your software's documentation to determine how to perform this fix.

Problem 2: Heteroskedasticity The ordinary-least-squares regression model also assumes that though the errors are independent from one another, they have constant variance. This implies that large values of a dependent variable are no more likely to have large errors than are small values of the dependent value. Though this is a statistical requirement, there may well be economic reasons why it fails to occur. Kennedy (2003) points out that when you are considering the relationship between consumer spending and income, it would not be unlikely for you to discover that the "errors" in spending increase as income does.

The consequences of heteroskedasticity are similar to those of autocorrelation: the standard errors on the estimates of the coefficients are biased. As a result, one cannot draw correct inferences about the statistical significance of the parameter estimates.¹⁰

NOTES FOR NOVICE RESEARCHERS

Assessing the Possibility of Autocorrelation

The regression results presented in Figure 11.4 give the estimated Durbin-Watson statistic as 1.43. For a sample size of 45 and two explanatory variables the lower limit for the critical DW is 1.43, while the upper limit is 1.62. Since the estimated Durbin-Watson statistic is near the bottom of the ambiguous range, this strongly suggests that autocorrelation is a problem. To address the problem, we reestimate the model using a first-order autoregressive term. The results are as follows:

Dependent Variable: CONS

Sample: 1990:4 2002:1

Included observations: 46

Variable	Coefficient	Std. Error	t Statistic	Prob.
C	−321.68	288.51	−1.156	0.2711
W5000	0.025	0.0105	2.391	0.0215
REAL_DPI	0.942	0.0617	15.257	0.0000
AR(1)	0.473	0.1525	3.100	0.0035
R-Squared:	0.9943			
Adjusted R-Squared:	0.9939			
Durbin-Watson:	2.2277			
F:	2394.7			

Note that the Wilshire 5000 is now statistically significant, and the marginal propensity to consume out of income (.942) is now less 1.00, which is consistent with theory.

The easiest way to detect heteroskedasticity is to examine a plot of the errors against each explanatory variable. If the errors are roughly the same as the explanatory variable increases, there is probably no heteroskedasticity. If the errors appear to increase or decrease in magnitude as the explanatory variable increases, you may have a problem. You may then want to go on to a more sophisticated test, such as the Goldfeld-Quandt test or the White test.

Most econometric software allows the user to correct for heteroskedasticity. As a result, when the model is estimated the *t* scores will be correctly

reported and the coefficient estimates correctly assessed. Consult your software's documentation to determine how to perform this fix.

Problem 3: Simultaneous Equations Bias The ordinary-least-squares regression model assumes that all the explanatory variables are truly independent or exogenous. That means that the explanatory variables are determined outside the model being estimated. Consider, for example, the demand for pizza on the part of a plumber named Susan. We generally assume that an individual's demand for pizza depends on his or her income. Now, Susan's income is determined by the wage rate she earns as a plumber multiplied by the number of hours she works. As a result, Susan's income is independent of her purchases of pizza. Or to put it a different way, her income is exogenous, and we could estimate her demand for pizza as a function of her income without difficulty.

But suppose that, instead of being a plumber, Susan manages a pizza parlor and she works on commission. Thus, in contrast to the previous case, her income would be affected by the number of pizzas she buys (as well as the number that she sells). In other words, her income is **endogenous**, or determined inside the model being studied. In this case, if we estimated her demand for pizza, the coefficient on income would be biased.

When we estimate Susan's demand for pizza, we want to determine the effects that income has on her pizza purchases. The problem here is that Susan's income is also a function of her spending on pizza, so that an increase in her income leads to an increase in her spending on pizza, which leads to an increase in her income and another increase in her spending. As a result, the coefficient showing the effect of the effect of income on demand will be overestimated.

Now this example is an extreme one by design, but the problem is a general one. When an explanatory variable is endogenous, the resulting parameter estimates will be biased. In fact, this is called **simultaneous equations bias**.

The solution is to take into account the relationship between the dependent variable (e.g., Susan's pizza purchases) and the explanatory variable (e.g., Susan's income) by using a more sophisticated estimation technique than ordinary least squares. Such a technique is beyond the scope of this book, but the problem of possible simultaneous equations bias is nonetheless something all researchers should be aware of.

Problem 4: Specification Error To obtain valid estimates of regression coefficients, we must specify the model being tested correctly. There are two aspects to model specification. First, the model must include the correct explanatory variables; the reason for this should be self-explanatory.

Second, the hypothesized equation must have the correct functional form. In other words, if you are using ordinary-least-squares regression, the relationship being estimated must be linear, or approximately linear.

This means that one can commit three kinds of specification errors. First, the model may be missing one or more relevant explanatory variables. Second, the model may include one or more irrelevant explanatory variables. Third, the underlying relationship may not be a linear one.

If you leave out a relevant explanatory variable, your estimated coefficients will be biased, and the estimated t statistics will be artificially reduced. This will make it appear that your coefficients are less statistically significant than they are. The effects of omitted variables show up either through the constant term or in the error term. Thus, a large and statistically significant constant term suggests that the model is incompletely specified. Similarly, autocorrelation may indicate an omitted variable. This is why theorizing before estimation is so important. It also suggests a solution: returning to the theoretical drawing board and considering what other explanatory variables may be important to the model. Another option is to review the literature for more complete models.

On the other hand, if you include an irrelevant variable, the resulting estimated coefficients will be unbiased. However, they will have inflated standard errors and reduced t statistics, which again makes it harder to draw statistical inferences.

Finally, the solution to nonlinear models is also theory. Does theory suggest a linear or nonlinear relationship? If the latter, there are several ways to proceed. One way is to use OLS and hope the relationship is approximately linear. Another way is to transform the data into a linear form, as discussed in Appendix 11A. The third alternative is to use a nonlinear estimation technique. Such a technique is beyond the scope of this book.

Problem 5: Multicollinearity Multicollinearity occurs when two or more explanatory variables are highly correlated. Kennedy (2003) defines multicollinearity more formally as occurring if there exists an approximate linear combination of two or more variables. An easy way to check for multicollinearity is to look at the correlation matrix for the explanatory variables. A standard rule of thumb is that if the correlation between two explanatory variables is more than 0.8, multicollinearity is likely to be a problem.

The consequence of multicollinearity is that the regression software doesn't know how to allocate the joint effects of the two collinear variables on the dependent variable. So, as Kennedy (2003) points out, the software throws out the joint effects and only gives "credit" to the unique effects of

NOTES FOR NOVICE RESEARCHERS

Symptoms of Multicollinearity

One sign of multicollinearity is that small changes in the data sample cause significant changes in the parameter estimates. Another is a regression that yields high R -squared but low t statistics, so few coefficients are significant.

each explanatory variable. As a result, the parameter estimates are inefficient with high variances—as if the independent variable lacked sufficient variation. The parameter estimates are still unbiased, but the variances are inflated. Thus, the estimates are imprecise, which leads to low t statistics, making hypothesis testing difficult.

The solution to multicollinearity is to provide additional information, either in the form of *a priori* conditions (based on economic theory) or in the form of additional data. Another solution is to use an alternative form of an explanatory variable, one that hopefully is not so correlated with the other explanatory variable. For example, one could create a proportion, such as investment spending as a share of GDP, instead of a level, such as investment spending by itself.

Note that you may solve the multicollinearity problem by dropping a variable that is highly correlated with another in the model. However, this simultaneously creates a specification error, so you should avoid this "solution."¹¹

NOTES FOR NOVICE RESEARCHERS

A Closing Comment on Regression Problems

Sometimes novice researchers make too big a deal out of econometrics problems such as these. They think that if you can't correct the problem, then no valuable information can be learned from the regression. However, the correct interpretation is a lot more subtle than this. It is not that these econometric problems render the empirical results invalid; it just means that they are possibly less valid than they would be absent the problem. Thus, the researcher must exercise care and judgment in evaluating conclusions.

NOTES FOR NOVICE RESEARCHERS

Interpreting Sample Regression Results

We presented results for our sample regression, corrected for autocorrelation, on page 220. Our hypothesis was that stock market wealth has a positive impact on consumer spending. The estimated coefficient for stock market wealth was 0.02 and was statistically significant. Modigliani's seminal study estimated the coefficient to be approximately 0.06, but more recent studies have indicated that the propensity to spend out of wealth has fallen, so our estimate is plausible. The estimated coefficient for income is 0.94 and is statistically significant. As a large fraction, this coefficient is plausible on theoretical grounds. The adjusted R-Squared is above 0.99, and the F statistic, at 2395, is far beyond the critical F, so the model fits the data very well. On balance, the results are quite favorable; they do support the hypothesis.

Step 5: Evaluate the Test Results

The last step in the regression process is to evaluate the results. This is really the point of regression analysis. You have developed a hypothesis. You wish to assess the evidence in the real world to see if it supports the hypothesis or not. You have collected an appropriate and adequate data set, and deduced the predictions that would confirm your hypothesis. You run your regression to estimate the model. You have corrected all the regression problems you can, as described in Step 4. You then interpret the results as described in Step 3, that is, you ask "What do the results say?" Now you must answer the more difficult question: what do the results *mean* with respect to your hypothesis?

More often than not, the results will be mixed. Some estimated coefficients will match your predictions, and some will not. Some estimated coefficients will be statistically significant, and some will not. Remember that if an estimated coefficient is not statistically significant, then you should treat the coefficient as zero. Some will be economically significant, and some will not. Some will be plausible, and some will not. The overall model may fit the data well or it may not. All misses are not equal, however. Did the key coefficients, those that directly assess your hypothesis, match? If so, then you have at least some evidence that supports the hypothesis.

Ultimately, interpreting regression results is more of an art than a science; it requires judgment, not merely computation.

SUMMARY

Consider the hypothesis that Y is determined by some causes X_1 , X_2 , and X_3 . This hypothesis can be expressed as an equation: $Y = a + b_1X_1 + b_2X_2 + b_3X_3$. The coefficient b_1 shows the relationship between X_1 and Y , since if X_1 changes by one unit, Y will change by b_1 units. Similarly, the coefficient b_2 shows the relationship between X_2 and Y , and the coefficient b_3 shows the relationship between X_3 and Y .

The coefficient a shows everything besides X_1 , X_2 , and X_3 that affects Y . In order to test the hypothesis, we need to analyze data to estimate the values of b_1 , b_2 , b_3 , and a . This is the role of regression analysis. In other words, when conducting a regression analysis, we are given X and Y , and the objective is to find values for the b 's and a in the equation.¹²

- Regression analysis can be broken down into five steps. Note that we will present them here in a slightly different order than we did in the text.
- The first step is to state the hypothesis of your research. What is the dependent variable? What are the explanatory variables? What are the expected relationships between each explanatory variable and the dependent variable? In other words, what coefficients will confirm the hypothesis?
- The next step is to test your hypothesis by performing the regression using appropriate software.
- Once you obtain results, you should check to see if the regression suffers from common problems, including autocorrelation, heteroskedasticity, simultaneous equations bias, specification error, or multicollinearity. Note that at least several of these problems can be addressed before the estimation—for example, by carefully thinking about model specification or by examining the correlation between explanatory variables. To the extent possible, fix any problems you discover.
- Once you have the most correct results you can obtain, you are in a position to interpret the results. In other words, you need to answer the question: What do the results say? Are the estimated coefficients statistically significant? Are they economically significant? Are they plausible in the light of theory and previous studies? How good a fit is the overall regression model? Finally, how closely do the estimated coefficients match the expected coefficients?
- The last step in a regression analysis is to evaluate the test results in order to draw conclusions about your hypothesis. In other words, what do the regression results mean for your hypothesis? On balance do they confirm or reject the hypothesis?

NOTES

1. Note that, strictly speaking, we mean linear in parameters rather than linear in variables. See, for example, Appendix 11A, which explains how some equations that have nonlinear variables can be transformed so as to be amenable to linear regression.
2. The issue of linear versus nonlinear regression is part of the more general issue of model specification. Model specification involves selecting a regression model that accurately characterizes the underlying relationship between the dependent and explanatory variables. There are three issues here: including all relevant explanatory variables, excluding any irrelevant variables, and using the correct functional form (i.e., linear or nonlinear) for the equation.
3. This is equivalent to the bivariate model $C = a + b \text{SMW} + e$. Alternatively, we could have plotted C and any other explanatory variable in Equation 4.
4. Conventional assumptions include: (1) that the errors are normally and independently distributed with mean zero and (2) that the errors have constant variance.
5. You should confirm this result, that $t = b/SE$, with actual regression results.
6. 1.68 is the critical t for 40 degrees of freedom; for infinite degrees of freedom, the critical t is 1.64.
7. The standard error is the same one in Figure 11.5. It is implied by the data in Figure 11.5 and is usually reported by the statistical software.
8. By estimating a demand or supply function in log-linear form, the resulting coefficient estimates are the respective elasticities. For more on the use of log linear functional forms, see Appendix 11A.
9. Econometricians use the term *best estimate* for this ideal. The precise meaning of *best* goes beyond the scope of this text. For an explanation, consult any econometrics text.
10. More formally speaking, both autocorrelation and heteroskedasticity cause the coefficient estimates to be unbiased but inefficient.
11. This assumes that the collinear variable truly belongs in the equation according to economic theory.
12. By contrast, the process of simulation or forecasting takes values for a and the b 's as given. Then, given values for X , one can predict corresponding values for Y .

SUGGESTIONS FOR FURTHER READING

- Gujarati (2003)*—Very pragmatic approach to econometrics. Less a study of the field of econometrics than of how to *use* econometrics in research.
- Kennedy (2003)*—Excellent guide to econometrics; more intuitive than most texts. This is one of the books graduate students have bought for years to supplement their assigned econometrics texts.
- McLagan (1973)*—Readable introduction to regression analysis for nonexperts published in the Federal Reserve Bank of Philadelphia journal *Business Economics*.
- Studenmund (2001)*—Very intuitive undergraduate econometrics text.
- Wyryck (1994)*—Nicely done introduction to econometrics for undergraduates, especially in its discussion of dummy variables and qualitative dependent variables.

EXERCISES

1. Suppose you plan to estimate a marginal cost function where $MC = f(w, r, Q)$. What would you expect the signs on the estimated coefficients to be if the cost function is valid?
2. Consider the Cobb-Douglas production function for the U.S. economy over the time period 1980 to 2000: $Q = TL^a K^b$. If we take the natural logarithms of both sides of the equation, we obtain a log-linear production function: $\ln Q = \ln T + a \ln L + b \ln K$. What do you expect the signs to be on the estimated coefficients of this production function and why?
3. Collect thirty years of annual data for U.S. GDP, employment, and capital stock. Take the natural logarithms of each variable and estimate the log-linear production function. Interpret the results of the regression. Are the estimated coefficients statistically significant or not? Are the coefficients economically significant or not? Are the coefficients plausible? To what extent do the estimated coefficients meet your expectations?
4. How good a fit is the overall regression model?
5. What does the Durbin-Watson statistic suggest regarding the possibility of autocorrelation?
6. Examine the correlation matrix for the three explanatory variables. Is there any indication of multicollinearity?
7. Is there any evidence of other common regression problems?

Data Transformations for Regression Analysis

Sometimes it is necessary to transform your data so you can either put the data in the form required for your model or to solve an econometric problem. Natural logarithms are particularly useful in this respect. Consider the Cobb-Douglas production function, which is the most widely used version of an aggregate production function:

$$Q = TL^aK^b \quad (16)$$

where Q is output, T is an index of technology, L is the labor supply, K is the capital stock, and a and b are parameters related to the marginal products of labor and capital. This function is nonlinear, so we can't estimate it directly using ordinary least squares. If, however, we take the natural logarithms of both sides of the equation, we obtain a log-linear production function: $\ln Q = \ln T + a \ln L + b \ln K$. Notice that this is like any other linear function: $Y = a + b_1X_1 + b_2X_2$. Therefore, estimating by using the ordinary-least-squares model is straightforward. Notice also that, with this form, the coefficients are estimates of the elasticities of output with respect to labor and capital. Similarly, if you estimate a log-linear demand or supply function, $\ln Q_d = \ln C + a \ln P + b \ln Y$, the coefficients are the estimated elasticities. For example, in this case, a is the price elasticity of demand, and b is the income elasticity of demand.

Suppose theory suggests that your model is nonlinear, but that taking logarithms will not help. The classic example is a wage function in which wages are hypothesized to increase as a function of age, but at a decreasing rate:

$$\text{WAGE} = a + b \text{AGE} + c \text{AGE}^2 \quad (17)$$

where $b > 0$ and $c < 0$. To estimate the function, you first need to create a **squared term** for age, using a spreadsheet or your statistical software: $\text{AGE_SQUARED} = \text{AGE} * \text{AGE}$. Then you can run the regression using WAGE , AGE , and AGE_SQUARED .

It is common to use a **time trend** in regression analysis—for example, to serve as a proxy for technology. The assumption is that technology improves by a constant amount each year. To create a time trend, you simply create a variable that increases over the sample and thus has values of 1, 2, 3, and so on.

Estimation Using Qualitative Variables

Up to this point, we have been assuming that all variables we are using for empirical testing are quantitative in nature, that is, they are variables with a wide and more or less continuous range of values, such as incomes, prices, ages, and so on. For many applications, however, you may need to employ qualitative variables. Qualitative variables come in two varieties: dummy variables and limited numerical variables. Dummy variables are those defined by nonnumerical values, such as male versus female, college graduate or not, and racial or ethnic background—for example, African American, white, Hispanic, Asian American, Native American. Limited numerical variables are those whose values take only a limited range of integer values, such as the number of children in a family. Qualitative variables can be used either as independent or dependent variables.

Qualitative Independent Variables

When the qualitative variables are on the righthand side of the equation, that is, when they are independent or explanatory variables, the solution is to use a **dummy variable** approach. The idea behind dummy variables is that there is a discrete effect as a result of gender, or education, or race or whatever the qualitative variable is.

The dummy variable is constructed so as to have a value of one or zero, depending on the case chosen. For example, you might use a dummy variable for gender in examining the wage gap. By letting female have the value of one and male have the value of zero, if the estimated coefficient showed a negative sign it would indicate how much lower female wages are than male, when holding all other factors constant.

Dummy variables can also be used for multiple categories, such as ranges of education: high school dropout, high school graduate, person with some college, college graduate. The technique is the same. The dummy variable for the dropout category would have a value of one for those individuals who did not complete high school and a value of zero for everyone else. The dummy variable for the graduate category would have a value of one for those individuals whose highest level of education was high school and a value of zero for everyone else. It is important that one dummy variable be omitted from the estimation; otherwise, the regression

will exhibit perfect collinearity between the dummy variables (and in all likelihood the estimation will blow up). (That's why we only use one dummy for the bivariate choices like male/female.)

Dummy variables are also used in time-series studies of policy analysis, where the two conditions are "before the policy was implemented" and "after." Another example of dummy variables is the use of **interaction terms**. Sometimes explanatory variables affect the dependent variable in a more subtle way than we typically assume. For example, we know that wages are influenced by both work experience and gender. Suppose we are interested in seeing if the returns to experience are affected by gender. We could define work experience (EXP) as number of years working. Then we could define a dummy variable for gender, MALE, where Male has a value of one, and Female has a value of zero. Finally, we could define an interaction variable $\text{MALE_EXP} = \text{MALE} * \text{EXP}$. This yields the following model to be estimated:

$$\text{WAGE} = a + b \text{EXP} + c \text{MALE} + d \text{MALE_EXP} \quad (18)$$

Thus, if you are male, your returns to experience are $b + d$, while if you are female, your returns to experience are b . If the estimated value of d is statistically greater than zero, then we conclude that males are rewarded significantly more than females for equivalent increases in experience.

The use of dummy variables causes no problems with ordinary least squares regression. You estimate your model in the same way as you would using quantitative variables.

Qualitative or Limited Dependent Variables

For some research, it makes sense to use a dummy variable on the left-hand side of the equation—that is, a dummy or limited dependent variable. Suppose you wished to model the determinants of housing choice by college students: on-campus or off-campus. The choices are binary: either they live on-campus or off. Recalling that dummy variables have a value of zero or one, estimating such an equation using ordinary least squares runs into several problems. If the estimated equation predicts a value that is in between zero and one, what does that mean? Similarly, what does it mean if it predicts a value less than zero or greater than one?

The solution is to use an estimation procedure that forces the dependent value to lie between zero and one. There are two such commonly used methods, the logit model and the probit model. For details on how to use these methods, consult any econometrics text.

Communicating the Results of a Research Project

"Science is an instance of writing with intent, the intent to persuade other scientists, such as economic scientists."

D. McCLOSKEY

A research project is not complete until its findings are communicated to a larger audience, namely, the scientific community. This is how knowledge progresses. This communication is performed in three steps. First, the researcher drafts a written report summarizing the results of his or her research. Second, the researcher presents the report orally at professional workshops or conferences. Finally, the researcher submits the research report for publication in the form of an article or book.

Note that each of these three steps may be done more than once. The report goes through multiple drafts as the author attempts to perfect his or her argument. The primary purpose of giving a research presentation is to gain feedback from experts in the field. After the presentation the researcher is likely to revise the research report again, based on the questions and comments he or she received. The researcher may then give the revised report in another research presentation. It is common for papers submitted for publication to be rejected. The author does not give up, however; rather, he or she looks for a new journal or new publisher until publication is assured.

This chapter explains the nut and bolts of this research communication process. The first half spells out how to write a research report, and the

NOTES FOR NOVICE RESEARCHERS

Venues for Undergraduate Research

A growing number of conferences offer opportunities for undergraduate economics students to present their research.

Examples of these include:

- the annual meetings of the Virginia Social Science Association (<http://www.vssa.net>)
- the Eastern Economic Association (<http://www.iona.edu/eea/>), and
- the National Conference on Undergraduate Research (<http://www.ncur.org/>).

Additionally, several journals exist that publish undergraduate economics research. Among these are the following:

- *Issues in Political Economy* (<http://www.elon.edu/ipe>)
- *University Avenue Undergraduate Journal of Economics* (<http://www.econ.ilstu.edu/UAUJE/default.html>)
- *The Student Economic Review* (<http://econserv2.bess.tcd.ie/SER/>)
- *SOAS Economic Digest* (<http://www.soas.ac.uk/SED/home.html>)
- *The Visible Hand* (<http://www.rso.cornell.edu/ces/publications.html>) and
- *Opus 1* (<http://www.opus1.org>)

second half describes how to prepare and give a research presentation. Since the report summarizes the process the researcher has undertaken, we will make numerous references back to earlier chapters.

Writing the Research Report

In Chapter 5, we introduced the idea that writing is a product of an intellectual endeavor. In this chapter, we present an in-depth example of that process: how the written research report becomes an outcome of the research process. Research reports come in several varieties. They include research papers (both undergraduate and graduate), under-

Figure 12.1 Components of an Empirical Research Paper in Economics

Title
[Abstract]
[Table of Contents]
[Acknowledgments]
Introduction and Literature Survey
Theoretical Analysis
Empirical Testing
Conclusions
References
[Data and Other Appendices]

graduate and masters' theses, doctoral dissertations, scholarly journal articles, monographs, and books. What we present here applies to all types of research reports, though what comprises a section of a research paper or journal article likely takes up a chapter or more in a book-length report. Journal articles tend to be the most condensed. Theses and dissertations tend to be longer since the author needs to demonstrate his or her command of material that would be assumed in a journal article or even a book.

The purpose of the written report is to present the results of your research, but more importantly to provide a persuasive argument to readers of what you have found. Recall from our earlier study of critical reading (Chapter 6) that there are two ways to think about the form of scholarly work. The first is format, and the second is argument. Format is more explicit, but argument is more important. The purpose of research is to advance knowledge in a field by providing a convincing argument supported by logic and empirical evidence. That is what your research report should be.

Your paper should be organized according to the format shown in Figure 12.1, which we first introduced in Chapter 6. The optional parts of the report are listed in brackets. Feel free to use section headings, as they will make the paper more readable.

As you draft your paper, keep in mind that the objective is to build a convincing case for your conclusions, namely, that the hypothesis is confirmed (hopefully!) or rejected based on the logic of your analysis and your empirical evidence. You may find it useful to review Chapter 4 on how to construct arguments.

Introduction

The purpose of the introduction to the research report is to provide the rationale for the research. This rationale should address four issues: What is the nature of the issue or problem the research investigates? Why is it worthy of investigation? What have previous researchers discovered about this issue or problem? What does your research attempt to prove? In other words, what is the contribution that your research will make to the literature?

The introduction is a critical part of the research report because it is often here that readers decide whether or not to continue reading the rest of the report. As the author you need to think of a way to persuade the reader to continue.

Start the introduction by sketching out the problem that the research addresses. What evidence can you offer to describe the issue? Why is it a problem? Is it a public policy issue, a social problem? Is it purely an intellectual puzzle? Who would be interested in the problem? It is often helpful to cite appropriate statistics to illustrate the magnitude of the problem or, better yet, use a figure or chart. The following paragraphs, which begin the introduction of a typical research paper, provide an admirable example of how to frame the problem:

Researchers are unsure of when and where the human immunodeficiency virus that causes the Acquired Immune Deficiency Syndrome originated. What is certain, though, is that HIV/AIDS has become a pandemic disease since its widespread recognition as a major global health crisis in the late 1980s and early 1990s.

This disease represents a severe development crisis in sub-Saharan Africa, the most afflicted region in the world. More than 29.4 million people are living with HIV/AIDS in sub-Saharan Africa, including 3.5 million new infections in 2002. Sadly, an estimated 2.4 million Africans died from complications with HIV/AIDS in the past year. National adult HIV prevalence has reached astronomical proportions, exceeding 30% of the adult population in four southern African countries: Botswana, Zimbabwe, Swaziland, and Lesotho. Even South Africa, widely considered the most developed country in Africa, has been touched by the far-reaching socioeconomic consequences of HIV/AIDS. It has become the most afflicted country in the world, with 4.9 million people living with HIV/AIDS.

Though it is not yet clear where this introduction is going, the reader should have no doubt that the HIV/AIDS epidemic in southern Africa is a substantial social problem.

NOTES FOR NOVICE RESEARCHERS

Using Figures Effectively

The purpose of charts, tables, and other graphics is to summarize and illustrate the argument in the text. Every figure should be designed to be easily understood independent of the text. Each figure should have a clear explanatory title. Each title should include two parts: a reference number (e.g., "Figure 12.1") and a description of the figure (e.g., "Components of an Empirical Research Paper in Economics"). The description should be as precise as possible, rather than vague. If the figure shows a graph, axes should be clearly labeled, with the units clearly identified. If more than one concept is being graphed, make sure the reader can clearly tell which is which (e.g., by using diamond symbols, dots, and/or x's for each element). If the figure shows a table, the numerical values in the columns should be aligned with the decimal points. In addition, each numerical value should have the same number of digits to the right of the decimal point.

Figures are best located in the text rather than as appendices. An exception is if you are running numerous regressions. In this case, you should summarize your results with a table in the text, but put the complete regression results in an appendix.

Make sure you refer the reader to each figure in the text of the report. Otherwise, they may not read the figure, which means it is not serving its purpose. This is especially true for figures that are relegated to appendices.

Next, explain to the reader why the topic is interesting and significant. We discussed what makes a research question interesting and significant in Chapter 2. Is it, for example, a hot topic among experts in the field? A research question is interesting if it addresses a problem that hasn't been entirely sorted out by the experts; in other words, some aspects of the problem have yet to be understood. A research question is significant if it deals with a substantial problem, one that has major impacts on the field.

The next step in the introduction is to summarize what has been done already to study the problem. This is the literature survey, which will be discussed in the next section of this chapter. Use that as a springboard to explain what your research is designed to accomplish.

NOTES FOR NOVICE RESEARCHERS

Write the Introduction Last!

Many experts think that you should write the introduction to your paper last rather than first. This makes sense since you want to hook your reader with the introduction, so you'll want to know what your paper says before you write its beginning. Alternatively, you may draft the introduction first but revise it last. Note also that if you have written a research proposal, much of the proposal itself can form the basis for the introduction to your report.

The Written Literature Review

An important component of a research report is the literature review or survey. We discussed the process of reviewing the literature in Chapter 3. In a thesis or longer work, the written literature review may be a part or chapter separate from the introduction, but in research papers or journal articles it is usually included as part of the introduction.

A literature review is a summary of the major studies that have been published on a research topic. The literature review should accomplish three goals. First, it should identify the major findings on a topic up to the present. Second, it should point out the principal deficiencies of these studies or provide a sense of what is lacking in the literature. Last, it should conclude by leading into your research question, by explaining how your research proposes to contribute to the literature or address some shortcoming of a previous study. Thus, the primary purpose of the literature survey in a written report is to provide a justification or rationale for your research.

Writing the Literature Review

Novice researchers tend to write excessively long literature surveys. This is understandable since they are likely to have a lot of notes on the studies they have read but perhaps fewer notes on and less confidence in their own original work. After all, summary is easier than analysis. It is important, however, not to allow the literature survey to "*dominate your paper*," as Booth et al. (1995, 163) put it. For most research papers, the literature survey should be brief.

Students sometimes ask how many sources should be included in the literature survey. The answer depends on how many major studies have been

NOTES FOR NOVICE RESEARCHERS

What a Literature Survey Is and Is Not

Novice researchers sometimes seem confused about what exactly should be included in a literature review. The literature survey is not

- a list of potential sources of information about your topic, or
- a list of sources that you reviewed, or even
- a list of summaries of the sources you reviewed.

Rather, the literature survey should be a summary of the major research done on the topic you are working with, where the objective is to give the reader only the background that will be necessary for them to understand your research. In other words, you shouldn't summarize the studies you discuss in their entirety; rather, you should only highlight the parts that shed light on your research question. Nor should you discuss studies you read but decided weren't directly related to your research. Note also that the literature survey should not include popular publications, only scholarly ones.

completed on the topic. If you only report one or two sources, readers may suspect that you haven't put enough effort into searching the literature. You don't want to miss a major study, since at best it will make you look careless and at worst it may weaken the rationale for your research.

If you have cast your net wide enough to find all the significant studies pertaining to your research question, you will undoubtedly read studies that you discover aren't really germane to your topic. These should be omitted from both the literature survey and the reference list at the end of the paper. Indeed, if you haven't found irrelevant studies you probably haven't surveyed widely enough. If you have doubts about whether or not to include a source, see if you can write a brief one-line description of its relevance to your study. If you have trouble doing this, you should probably leave it out. Never include in your report discussion of any source that provided only background information.

For each study that you include in your survey, you should give complete citation information. You should use the appropriate style for citations in the text of your survey (e.g., "Anderson (2003)" and in the references. (If you need to refresh your memory on citation styles, review Appendix 3A in Chapter 3.) You should also indicate the question that the

author examined, the author's findings, and anything about the methodology he or she used that might be appropriate to your study—for example, data sources or the type of empirical testing employed. Remember to focus your commentary on those aspects that shed light on your research question:

Armstrong (1991) was able to accurately predict the incredible socioeconomic consequences of the HIV/AIDS epidemic, theorizing that HIV/AIDS would alter the quality and size of the labor force and impede human capital formation. As soon as adequate data became available, formal data analyses were conducted. Over (1992) found the net effect of the AIDS epidemic on the growth of Gross Domestic Product (GDP) per capita to be a reduction of a third of a percentage point in the ten most afflicted sub-Saharan African countries. Bloom and Mahal (1997) concluded the opposite, reporting that the impact of AIDS was statistically insignificant, by using a data analysis examining the effects of AIDS on GDP growth per capita in 51 industrial and developing countries. Still, recent reports such as the World Health Organization (WHO) report by the Commission on Macroeconomics and Health predict a decline in economic growth due to the HIV/AIDS epidemic.

Some studies can be summarized with a sentence. Others, for example, a paper whose model you are going to base your own research on, will need more detail. You want to be as concise as possible, while getting the important points across.

Booth et al. (1995) suggest that if a source is an important one for your research, you should name it directly in your summarizing sentence:

Ruhm (2002) has found that the adverse effect of parental work hours on children's cognitive development is statistically significant.

If the source is less important, however, cite it parenthetically at the end of the sentence:

Previous studies have found a mix of results, with little or no statistical significance. (Blau, 1999; Barling & MacEwen, 1988.)

You should conclude the literature survey by leading the reader back to your research project, explaining how your research will contribute to the literature:

Since few studies on the subject exist, research on the impact of HIV/AIDS on productivity and the labor supply must continue to be conducted using accurate and more recent data on the prevalence of HIV/AIDS. This study improves upon the existing research by using actual AIDS data collected by each country, rather than predictions,

in attempting to discover the extent of the impact of AIDS on the labor supply in Botswana, Zimbabwe, Swaziland, and Lesotho; the four most afflicted countries in sub-Saharan Africa.

Theoretical Analysis

The purpose of this part of the research report is to present the theoretical analysis of the issue or problem you are investigating. This is also described as presenting your theoretical model. We discussed how to develop this analysis in Chapter 7. Recall that the objective of research is to advance knowledge in a field by providing a convincing argument that is supported by logic and empirical evidence. This section of the research report presents the logical evidence as a deductive argument; the empirical evidence is presented in the next section of the report.

So, how does one explain the theoretical analysis? There are several ways to do this, so let's explain them from simplest to most complex, using examples from student papers. The basic approach is straightforward: you must clearly describe the theory you are applying to your research problem, explain in detail why it is relevant, and then sketch out how it diagnoses the problem. This diagnosis is the research hypothesis.

If you are applying a common economic theory, such as the theory of demand, you should explain how your problem is an application of that theory and what the major elements of your analysis are. What plays the role of price, income, tastes, prices of related goods, and the like in your analysis? Note that since you are applying a common theory you should assume that your readers will be familiar with it. There is no need to derive the law of demand or any other results that all students of economics know. Indeed, doing so will detract from your report.¹

In order to demonstrate that consumers prefer financing incentives, the theory of demand will be applied to the sale of new automobiles in the United States. According to the theory of demand, the demand for a certain good is a function of changes in the price of that good, the disposable income of consumers, and consumers' preferences. . . . Because a change in the interest rate on a loan used to pay for a good is effectively equal to a change in the price of a good, the two variables should have [a comparable] impact on the sale of new automobiles. . . . I hypothesize that the interest rate elasticity of demand for automobiles will be larger than the price elasticity of demand. . . . [I]n accordance with the theory of demand, I predict that the price and rate elasticit[ies] . . . will be negative and that the income and measure of consumer sentiment elasticit[ies] . . . will be positive. (Beck, 2003)

If you are modifying an existing applied theory, you should first restate that theory in enough detail so readers can fully understand what the earlier researcher did. You should not assume that they will be familiar with the previous theory. Next, you need to explain how your research is an application or extension of that applied theory. Often, you need only indicate that the earlier theory was from the literature on your problem and that you have made a potential improvement.

The author of the following quotation modifies a model by Ehrenberg et al. (2001), which she indicates provides “a solid framework for the studies of the executive compensation in higher education.” After explaining the model in some detail, and observing that its empirical results were weak, she brings in additional factors that she hopes will improve the model:

I use an empirical model similar to Ehrenberg et al. (2001) to analyze the factors influencing the salaries of liberal arts college presidents. In addition to college performance variables and human capital variables [used by Ehrenberg et al.], I include social capital variables that trustees may value in a president—leadership skills, social capital, academic and managerial reputation. (Sorokina, 2003)

As you sketch out the logic of your theoretical analysis it is helpful to remember that you are trying to develop a deductive argument that culminates in the research hypothesis, which is often expressed in the form of an equation of the model. Recall that deductive reasoning starts from one or more assumptions and derives a specific prediction from them.

If your analysis includes formal mathematical reasoning, either an optimizing approach or ad hoc, explain it here. What are the assumptions from which you start your analysis? What mathematical manipulations did you perform on the assumptions in order to obtain your research hypothesis as the result?

In an empirical research paper, emphasis tends to be less on the theoretical analysis than on the empirical analysis. This is as it should be. At the same time, you still need to do a complete job of explaining your theoretical analysis. Even if you are merely applying a common economic theory or modifying an existing applied theory, the theoretical analysis is more important than you may think. This is because the validity of empirical testing *always* depends on the underlying theoretical model. Empirical testing without theory only shows correlation rather than the causation that researchers seek. The causal inferences must come from the theoretical analysis!

NOTES FOR NOVICE RESEARCHERS

The Importance of Explaining the Theoretical Analysis

The purpose of empirical testing is to validate your hypothesis. Most undergraduates can successfully apply statistical software to a data set to perform a statistical analysis. The underlying question for any statistical analysis, however, is what does it mean? The results of the statistical analysis have no economic meaning unless they are interpreted in the context of an economic theory. The theory always has to come first. The theory explains where the hypothesis came from. The hypothesis is a proposition about cause and effect. If the statistical results are consistent with the hypothesis, then and only then do they support a causal interpretation.

Empirical Testing of the Analysis

The purpose of the empirical testing part of the research report is to provide the empirical evidence for your research argument. The theme of this section of the paper can be summarized as: Given your hypothesis, how did you test it and what were your findings?

Your discussion should include a number of elements, including the data used, the empirical model and type of statistical analysis you employed, the results you hypothesized, your actual results, and your interpretation of the results. We discussed developing an empirical testing methodology in Chapters 10 and 11 and finding data and constructing a data set in Chapters 8 and 9. If you don't feel comfortable with this material yet, you might want to revisit those chapters.

You're not merely explaining what you did, but rather trying to persuade the reader that your approach was a fair and accurate test of the hypothesis. For that reason, you need to think about, if not explicitly address, questions such as: Does your empirical model adequately represent your theoretical one? Are the data good enough to accurately test your hypothesis? Does your statistical methodology discriminate well enough to get at the question you're asking and rule out alternatives? Does it adequately control for outside factors? Ultimately, you need to ask: if the empirical test obtains the best possible results, how confident can you be that the hypothesis is correct? Let's discuss each of these issues.

Testing Methodology Early in this part of the research report, it is helpful to have a clear statement of the testing methodology, including the type of statistical analysis employed and a brief description of the scope of the test:

This research uses a pooled Ordinary Least Squares regression analysis from 1990 to 1998 for Botswana, Zimbabwe, Swaziland, and Lesotho.

In order to test the hypothesis, this research performed an ordinary least squares regression using monthly data over the time period between 1996:1 and 2001:12.

You may wish to explain why the statistical methodology is an appropriate method for testing the hypothesis. Ordinary least squares may be considered the default statistical technique, though it may occasionally be adequate to use something less sophisticated like a statistical test of means. If you use any technique other than ordinary least squares, such as pooled OLS, generalized least squares, probit, and the like, you should explain why that is appropriate.

Empirical Model This section of the report should include an equation for the empirical model. The empirical model is the actual version of your theoretical model that you test empirically. (Note that if you use a statistical test that is less sophisticated than regression, there may be no empirical model per se.) As such, there should be a clear correspondence between the variables identified in the theoretical model and those employed in the empirical model. It is common that data for one or more variables in the theoretical model are not obtainable. Indeed, this is perhaps the principal difference between the theoretical model and the empirical model. There are two solutions to this difficulty. The first is to find data to serve as a proxy for the theoretical variable. The second is to omit the theoretical variable. Whatever approach you take should be explained. You should also explain how you obtained the empirical model from the theoretical model, in terms of functional form, mathematical manipulation, and so on.

The empirical model is this:

$$Q_D = a_0 + b_1 P_d + b_2 M + b_3 Y_D + b_4 A + e$$

where P_d is the average diamond price, M is the number of marriages, Y_D is the disposable income of the consumer, A is the event of advertising, and e is the error term.

Data Explain the sources and methods of the data used to test the hypothesis.² What is the data set, and where did you obtain it? What data

transformations, if any, did you perform? If this discussion is complex, briefly summarize it in this section, and then refer the reader to a more extension discussion in the data appendix. This section of the paper should be explained in enough detail to persuade the reader that the data are adequate to fairly test the hypothesis.

I compiled data from the Trade DataWeb at the U.S. International Trade Commission website on the quantity of diamonds imported, the average price of a diamond for each month, personal disposable income, and the number of marriages occurring each month. I also constructed a dummy variable to indicate when new advertising campaigns began. I found seventy-two data points for each variable, one for each month over a period of six years. This data is located in the appendix at the end of this paper.

Hypothesized Results Before you present your actual results, it is critical for the validity of your test to identify the results that are predicted by and thus would support your hypothesis. For example, you should state the algebraic sign and magnitude (if this is predicted by the theory) for each variable in a regression equation:

I expect certain relationships between these economic factors. I predict that as advertising (my measure of tastes and preferences) increases, demand for diamonds will also increase. Likewise, as consumers' personal disposable income increases, so will the demand for diamonds. There should also be a positive correlation between marriages and number of diamonds demanded. Finally, I predict that as the price of diamonds increases, the quantity of diamonds demanded will decline, although I do not expect this correspondence to be as strong as the other relationships.

Actual Results Display your empirical results in a table or figure that is clearly visible in the text. (Only if there are many regressions should you put the results in the appendix; even then, summarize the key results in a table in the text.) A sample is shown in Figure 12.2. The table should include the sample size, estimated coefficients for each explanatory variable, some measure of statistical significance (e.g., p value, t statistic, or standard error) for each variable, and goodness of fit information (e.g., Adjusted R -squared). It should also include any relevant test statistics for possible econometric problems (e.g., Durbin-Watson statistic). Like all figures, this one should be clearly labeled and titled. Finally, you should refer readers to the table, not simply assume or hope they will find it.

Interpretation of the Results Once you have presented the empirical results, you need to interpret them for the reader. We noted in Chapter 2

Figure 12.2 Regression Results

Dependent Variable: Q				
Method: Ordinary Least Squares				
Sample: 1996:01 2001:12				
Included observations: 72				
Variable	Coefficient	t Statistic	P Value	
C	-2881179.000	-5.976	0.000	
P	-2231.395	-7.726	0.000	
M	1.469	2.311	0.024	
DI	585.034	6.947	0.000	
ADS	-25958.720	-0.263	0.793	
	R-Squared	0.692	Durbin-Watson	2.385
	Adjusted R-Squared	0.668	F-Statistic	29.695

that this is the part of the research process that novice researchers seem to most overlook, and yet it is here that your research hypothesis stands or fails.

How do the estimated coefficients compare with the predictions of the hypothesis? Were the estimated coefficients statistically significant? Were they economically significant? How do the results compare with those of previous studies? What can you conclude about the results from the goodness-of-fit statistics? Were there any statistical problems that need to be corrected in order to obtain valid results? If so, make clear that these are the corrected results.

It is unlikely that you will obtain perfect results—for example, where every estimated coefficient has the correct sign and is economically and statistically significant. Discuss any anomalous results and suggest explanations, where possible.

The results of my regression were fairly strong. My R-squared value indicates that I explained approximately 69% of the data with my model. All of my variables were significant except for advertising. Price was highly statistically significant which I did not expect. I was under the impression that De Beers kept prices fairly constant, but this is apparently not the case. The price coefficient was negative, which agrees with the theory of demand, in that price has a negative effect on the quantity demanded. Marriages were positively correlated with diamond demand, which also agreed with my hypothesis. This variable was not as statistically strong as some of my other factors. I believe that this is due to the effect of substitutes; diamonds

NOTES FOR NOVICE RESEARCHERS

Presenting Statistical Results

Never cut and paste regression results from your statistical software into your paper. Rather, choose the results that you wish to present. It's good practice to identify the dependent variable in a regression as well as the sample size. Omit anything from the regression results that you don't understand, since you could be asked to explain it. (Hopefully, this doesn't include any of the important results described earlier.) Finally, don't show more than a few significant digits for most entries in the table.

are not the only stone presented in engagement rings, and not every bride receives an engagement ring. Disposable income was also a very strong variable. It is positively related to the demand for diamonds, which indicates that as consumers' disposable incomes increase, the quantity of diamonds demanded will increase as well. On balance, my regression results seem to support my hypothesis. The majority of my variables were significant, and the coefficients had the signs that I expected. The advertising variable was not very satisfactory. I was expecting it to be positively correlated with diamond demand, and highly significant. Unfortunately, the regression showed that it was in fact negatively correlated with diamond demand and not very significant at all. I believe that this may be the result of poor data. I could not find a definitive source for advertising data on the Internet, so I used the few press releases that I could find to construct the dummy variable. If I could run the regression again, I would like to find empirical data for advertising. Empirical data would result in a more scientific evaluation of the advertising variable. I would also like to find data on substitutes for diamonds, such as prices and quantities of sapphires. I believe that these additions will allow my model to explain a large proportion of the data, improving its accuracy.

Conclusions

The purpose of this part of the research report is to summarize your findings, that is, to restate your argument and conclude whether or not it is valid. You should note that conclusions are inductive: in light of the

NOTES FOR NOVICE RESEARCHERS

A Failed Hypothesis Does Not Mean a Failed Project

Note that it is perfectly appropriate to conclude that, based on the empirical testing, the hypothesis fails. This does not mean your research is a failure or that you will obtain a poor grade. Nor should you reverse your hypothesis so that it is confirmed by the testing—this is bad science.

statistical results, what can you infer about your hypothesis? To what extent did your empirical testing confirm your analysis?

If the results confirm the hypothesis, before you congratulate yourself recall that those results may also be consistent with alternative hypotheses! In other words, a weak empirical testing design may not rule out plausible alternatives. Based on your empirical tests and results can you rule any alternatives out? If not, you should admit this in your concluding remarks. By admitting to the weaknesses of your argument, you actually make it stronger.

If the testing did not confirm the hypothesis, can you suggest reasons why it didn't? Were the data or testing methodology flawed? What do you know now that you wish you had known when you began the project? What could be done to improve the results, given additional time and effort?

Finally, what can be concluded about the research question more broadly?³ What have you learned from the project that might help other researchers in the field? For example, the literature on the relation between parental employment hours and children's cognitive development lacks consensus about whether the effects are permanent or temporary. This suggests that something fundamental is missing from the studies to date. A recent study concluded by proposing that what is missing in past studies is controlling for the *quality* of child-care, data for which has not been available in the past. Thus, rather than continuing to test with existing data, a more promising approach could be to develop data for child-care quality.

Other Components of the Paper

Research reports include several other components, some of which are optional. These include a title, an abstract, a table of contents, acknowledgments, references, and data and other appendices.

Title All research reports should have a carefully composed title. A good title, like a good introduction, is designed to capture the reader's interest. One way to capture interest is to pose your research question in the title, especially if it raises an angle that hasn't been thought of before. An example might be *Is Divorce Indebting Our Children?* Another way is to highlight counterintuitive findings of your research.

A good title will have an enigmatic quality; it will identify the research topic but leave enough unsaid to draw the reader in. One way to do this is by developing a two-part title, where the first part raises the general topic, and the second part poses a more specific question. Examples of this include:

Intergenerational Mobility: Do First-generation Immigrants Lag Behind?

All Men Created Unequal: Trends and Factors of Income Inequality in the United States

Wage Discrimination in the NBA, Does It Exist for Foreign-Born Players?

Because online searches often used titles as targets, it is good practice to specify the key concepts of the research in your title. You should be particularly careful about avoiding vague titles that don't make clear what the paper is about, such as

Aggregate Reflections of Brand Loyalty or Unraveling the Size Effect

If the reader can't figure out what the paper is about, he or she is unlikely to decide to read it, however good the research may be.

It is common to begin a research report with a title page. In addition to the report's title, the title page should include the author's name and affiliation and the date of the report. For a term paper, the affiliation would list the course the research is for and possibly the instructor's name. For a thesis or journal article, it would indicate the author's academic department and institution.

Abstract It is good practice to write an abstract of your research report. Refer to Chapter 6 for how to do this. Remember that in this age of electronic databases the abstract is likely to be searched for keywords. The abstract should be placed either on the title page or with the first page of text.

Table of Contents Any book-length research report or thesis will benefit from a table of contents. Make this a separate page. A table of contents is not usually included in a journal article.

Acknowledgments It is considered good form to acknowledge any substantial assistance you received in a research project. Acknowledgments in an article are usually positioned as an initial footnote under the title. For

theses and other book-length reports, acknowledgments are put on a separate page or pages.

Reference List All scholarly works should include a reference list or bibliography. Indeed, the inclusion of references is one feature that distinguishes scholarly from popular publications. The reference list should provide complete bibliographic information for those sources that you cite in your paper, using the appropriate bibliographic style. Refer back to Appendix 3A in Chapter 3 for detailed information on bibliographic styles. In general, it is not good practice to include references for “background information” or any sources that you did not mention in the text. Such references are usually an indication of padding.

Data and Other Appendices Research reports occasionally include appendices. Appendices contain details that are not essential to grasp the re-

NOTES FOR NOVICE RESEARCHERS

Style Suggestions

What follows are suggestions regarding the style of a research report. Style is a matter of taste and common practice, so feel free to deviate from these if you have a good reason to.

It is common practice to put the abstract on the title page of the paper, after the title, author, and other information. The abstract should be single-spaced. Some authors use narrower margins for the abstract vis-à-vis the rest of the text, while still visually centering the abstract on the page. That's fine, but make sure you align the text of the abstract with the left margin or both left and right margins.

Don't center-align the text in an abstract.

It will make each line have a different length, which looks funny and unprofessional.

The body of the paper should be double-spaced. Footnotes and references should be single-spaced. Include page numbers in the document for all pages after the title page. Always refer in the text to any figures that you provided to support your argument; you shouldn't assume that readers will find and make sense of them on their own.

Before you begin drafting your paper, you may wish to review Chapter 5 on economic writing. Start writing the report early enough

that you will have sufficient time for revisions. You don't need to wait until your results are finalized to begin writing. It's not uncommon for researchers to draft the report, except for the results and conclusions, before completing their empirical work. You want to craft your argument to maximize its impact on your readers. For example, page-long paragraphs suggest insufficiently deep thinking about the story you are telling. Remember, a paragraph is supposed to explain a single thought.

Don't begin your literature survey by indicating that no previous research has been done on your exact topic. That is actually a quite normal occurrence. Rather, think of a way to lead into the literature that you have found useful in formulating your own research. Refer to the authors of a study as people; don't treat the study as if it were making the argument. Don't say “Francos and Schairer (2003) notes”; say “Francos and Schairer (2003) note.”

Don't discuss the data or other empirical issues in the theoretical section of your paper. Remember, the theory section is supposed to develop your hypothesis, not test it.

Always state the empirical results predicted by the theory before you present the empirical results. If you performed multiple iterations of the empirical test, minimize discussion of the *process* that you went through to obtain the results. Don't say: I did this; then I did this; then I corrected for that. Rather, summarize the problems you ran into and what you did to address them; then present and discuss the final results. Don't begin presenting your results by discussing R-squared, since this is not the most important aspect of testing hypotheses. Estimated coefficients, which are not statistically significant, are considered zero, so don't overemphasize their sign or magnitude. Similarly, don't discuss the estimated coefficients independently of their statistical significance. If you present your regression results in the form of an equation, be sure that under each estimated coefficient you include the standard error, *t* score or *p* value so the reader can assess the statistical significance of each estimated coefficient.

Don't change your hypothesis after obtaining your results. It doesn't make the paper stronger, and it violates the scientific method. Rather, stick with your original hypothesis, but report that it fails. The paper will be stronger, and you will get a better grade. (Failure of the hypothesis does not cause a lower grade! A low grade is given for a poorly argued paper.)

searcher's argument, but that some readers may wish to have easy access to. A data appendix includes all the data you used in the empirical testing of the research project, along with complete information about your sources and methods. Refer to Chapter 9 for more detailed discussion of how to construct a data appendix. Other appendices might include survey questions used to develop data or empirical results from multiple tests that would distract readers from your argument if their details were included in the text of the report.

Presenting Research Orally

The second product of a research project is an oral research presentation. There are different types of research presentations, depending on the audience and length of the presentation. Some are directed toward a general audience, either an audience of economists who are not experts on the topic or an audience of noneconomists. General presentations tend to last relatively longer, say, thirty to sixty minutes, to enable the presenter to develop enough context to make his or her remarks comprehensible.

At the other end of the spectrum are undergraduate honors, master's, or Ph.D. thesis presentations. These presentations are addressed to faculty who are assumed to know the discipline and the topic of the research. In general presentations, the speaker knows far more than the audience and needs to make his or her remarks general enough to be understood. In contrast, in thesis presentations the speaker is expected to demonstrate mastery of research techniques and the nuances of the topic. In a sense, it is as if the audience knows more about the subject than the speaker. Thesis presentations tend to last an hour or longer when questions and answers are included.

The majority of research presentations are those that scholars routinely engage in as part of their research. These presentations are between ten and twenty minutes long. They will be the focus of the remainder of the remarks in this chapter.

Though we tend to think of research presentations as providing researchers with an opportunity to disseminate research findings, more often their primary purpose is enable the researcher to obtain feedback from other scholars before submitting his or her work for publication. Journeyman researchers tend to fear criticism of their work, perhaps seeing it as criticism of themselves. By contrast, professional researchers look forward to criticism of their work because it helps them improve it. Though you might think an absence of criticism from other scholars means your research is flawless, it is much more likely to mean that no one has carefully considered your work! This is especially true of an initial pre-

NOTES FOR NOVICE RESEARCHERS

You Are the Expert on Your Topic

Despite what you may think, rarely will anyone know as much about your research as you do. The only possible exception is the faculty member who is supervising your research, and you can be sure that he or she will be on your side in a thesis presentation.

sensation. Like writing a draft report, essentially no one can present a research project perfectly on the first try. In this light, Missimer (1995) observes that, "If we can acquire the habit of seeing [research presentations] as opportunities to acquire wisdom, with no possibility of 'losing' an argument since we always 'win' insight, it will be easy to cultivate a spirit of welcome for any ideas" (p. 43).

An Oral Presentation Is Different from a Written One

Though both a research presentation and a research paper report the results of research, they are structured very differently. To begin with, you won't be able to cover as much material orally as you can in writing. At least, you won't be able to cover it well.

This is true for several reasons. First, as we noted earlier, you will have only a limited amount of time to present, often only about ten minutes. Second, it is substantially harder for an audience to assimilate an oral presentation than a written one. This is partly because of humans' ability to process aural information compared to visual information. Additionally, in written discourse the reader can always go back and reread. In an oral presentation, they cannot. This leads to the third reason.

An oral presentation should be given at a lower technical level than a written presentation.⁴ A research paper is implicitly addressed to an audience of experts in the field. Indeed, reviewers of papers submitted for publication are always chosen from such experts. By contrast, a research presentation should be addressed to a more general audience. Even if the audience includes economists who have backgrounds in the general area of your interest, few will be expert in your specific topic area. There is also a good chance that at least some in the audience are merely interested in the topic but have little or no previous expertise in it.

In short, a research presentation should *summarize* your research paper. It shouldn't attempt to replicate it in oral form. You should think of giving an oral presentation as telling a story; in other words, it should focus more

on the research problem or question and less on the technical details. Of course, some technical points will be necessary in the presentation, for example, the equations of the model and the empirical results. But any technical points or results should be illustrated with handouts or graphics to make them easier to understand. We will discuss each of these points in more detail in the following sections.

Preparing the Presentation

You may not realize it but anxiety over public speaking is very common, especially among those who lack experience with oral presentations. The key to minimizing such anxiety is to know your material. This involves two steps: planning what you will say and practicing the presentation until you know it well.

Let's explore how to plan an effective presentation, beginning with some general comments. First, your presentation should highlight what is novel about your research, whether it is a new or better model applied to the topic, a new data set, or a new statistical technique. Speech experts teach that an audience will only be able to take in and retain a limited number of ideas. As a result you must ask: what are the main points I want the audience to remember? This list should be relatively short. Remember, less is more.

Preparing your presentation "notes" is an intensely personal and subjective thing. It is often useful to make an outline of the major points, as well as include definitions of key terms or quotations, equations of the model, and empirical results. There is no single correct approach. Some speakers use detailed notes; others use more limited ones.

There is only one rule for presentations: Never read your remarks! Similarly, you probably shouldn't memorize your research presentation either. When reading or presenting a memorized report, speakers act as if the audience is not there. By contrast, what you want to do is engage the audience in a conversation. After all, you want their feedback. To do this, you want to "talk about" your research. You need to know the material, but you don't need to know each word.

The format for a research presentation is fairly well defined and similar to that of a research paper. A research presentation should be structured to have four parts: the introduction, the analysis, the results or evidence, and the conclusions.

When you prepare your presentation, you should think explicitly and carefully about the introduction. You begin by explaining what your topic is. It is often useful to demonstrate the issue or problem with a figure. McCuen et al. (1993) argue that the first minute or two of a presentation is the most

critical part because it is then that audiences decide consciously or unconsciously whether they will pay attention to the rest of the presentation.

So how do you engage the audience? By showing that the problem is interesting and worth exploring. What has been done in the past to investigate the topic? Why has past research not been entirely satisfactory? In other words, why is the topic still a problem for researchers? This is the part of the presentation that is most like storytelling. You should plan on spending a substantial proportion of your presentation on the introduction, say, perhaps a third.

The next two sections make up the body of the presentation, which should compose roughly half of the allotted time. You should plan to spend relatively more time on the section that discusses the principal contribution of your research, and relatively less on the other. The analysis section summarizes how your research addresses the deficiencies in previous work. Here you should describe the approach you took to analyze the research problem, presenting your model or methodology. The empirical section describes how you tested your analysis, the results you obtained, and the extent to which they confirmed your analysis.

You should also think explicitly about how you will conclude the presentation. The conclusion is almost as important as the introduction. It is here that you tell your audience what you want them to remember from the presentation. You can signal this by starting your conclusion with a phrase like: "Let me summarize . . ." or "In conclusion. . ." Such a statement acts like a road sign for the exit on a highway to your destination. It signals the audience that the presentation is almost over, and it naturally reengages their attention. After giving this signal you should state your conclusions clearly and concisely. In total, this section should take about 10 percent of the presentation's time.

Use of Visual Aids: Handouts, Transparencies, PowerPoint

We pointed out earlier that the main weakness of the spoken word as a means of conveying information is that visual comprehension is faster and more complete than aural comprehension. One solution to this very real problem is to use visual aids to complement your research presentation.

Visual aids include handouts, transparencies, or PowerPoint presentations. Handouts are printed pages, either distributed individually or as a package. Transparencies are slides that are displayed using an overhead projector. PowerPoint is computer software that creates presentations that can then be displayed as handouts, transparencies, or directly from the software using either a display panel, which sits on top of an overhead, or a separate data projector. Note that handouts and transparencies can also

NOTES FOR NOVICE RESEARCHERS

Opening Slide of a Presentation

The opening slide of a presentation should be like the title page of a research report. Put the title of your paper as well as your name and institutional affiliation on the opening slide. It's also common practice to include the name, location, and dates of the conference or workshop where you are presenting.

be created using word-processing software. Each visual aid format has pluses and minuses. Let's begin by discussing the use of visual aids in general; then we will examine the different formats.

Two types of information benefit from visual aids: information that you want listeners to remember and information of a technical nature. Think about an outline of your presentation. You might use visuals to show the titles of each of the major sections of the presentation or each of the main points in the presentation. You might use them to show definitions of key terms or quotations. You might also use them to list the conclusions of your presentation. Suppose you introduce the presentation with statistics or data to illustrate the research problem. Showing a figure or chart will make more of an impact on the audience than simply quoting the statistics. Remember, a picture is worth a thousand words. Additionally, you should always use visuals to show the equations of your model or to present empirical results. All tables, figures, and charts should be fully labeled so the audience can understand them without any instruction from you. However you use visual aids, the key rule is that visuals should be designed to *complement* the presentation—they should never become its focus.

PowerPoint can be a useful tool for organizing a presentation, whether in the form of handouts, transparencies, or directly via the software. Using your research paper or notes, you can “wing” an oral presentation, giving it without adequate thought or planning. You can't wing a PowerPoint presentation, since designing the presentation forces you to plan it.

When you design each slide, you are creating talking points. For slides to have their most powerful impact, they should not be crowded with information. You want the listeners to grasp each point with a minimum of

effort, and then return their attention to you, as you elaborate in more detail. Slides should show less information than you discuss, not more. The exception to this is statistical results where you may want to show all the results, but not discuss each explicitly. Note, however, that anything you include on a slide is grounds for discussion, so never cut and paste entire regression results. Rather, always create your own table that includes only the results you are willing and able to talk about.

Let the slides serve as the outline of your remarks rather than the text of them. Don't merely read your slides. The audience will read the slides faster than you can voice them, so if your presentation consists of reading the slides verbatim, you will lose your audience. The exception to this rule is slides that contain definitions of key terms or quotations. These it is appropriate to read.

As we observed earlier, each format of visual aid has pluses and minuses. Handouts are bulky to transport and take time to distribute. Transparencies can be cumbersome to use. The placement of the overhead projector can obstruct the view of the audience, especially when changing transparencies.

Whether using transparencies or PowerPoint, inexperienced presenters often fail to leave slides up long enough for audience members to read them, frustrating those who struggle to follow. Visual aids are supposed to enhance the presentation, not detract from it. This is especially problematic for technical material. Remember that while you are familiar with your statistical results, the audience is seeing them for the first time. It is good practice to direct attention to the screen where slides are displayed, but avoid speaking to the screen. Remember to face your audience. One way to make this appear natural is to put the laptop that is running the presentation on your podium, which you are standing behind. Then follow the presentation on the laptop rather than from the screen.

Handouts don't suffer from these problems, but they do transfer attention from you to your handouts. Know that as soon as you distribute the handouts the audience will begin reading them and not pay as much attention to what you saying. One way to alleviate this is to distribute handouts one figure or page at a time, right at the moment when you want to transfer attention to the figures.

The power of PowerPoint tends to take over the presentation—exactly what you don't want it to do. If you use PowerPoint, you should minimize the “bells and whistles.” Such features are like the annoying TV commercials in which the audience remembers the commercial but not the product it was advertising.

NOTES FOR NOVICE RESEARCHERS

Stay on Message!

Your research presentation should emphasize what is most important in the research. Novice researchers sometimes have a problem seeing the forest for the trees. They talk about all the details, but fail to highlight what is most important about the research. What is the hypothesis? To what extent did the empirical results confirm the hypothesis? Often the test of the hypothesis is the estimated coefficient of a single explanatory variable. When that is the case, be sure to highlight it rather than treating it like any other coefficient estimate.

Practicing the Presentation

Once you have prepared your presentation, it's a good idea to practice it at least a few times before the actual event. At least one of those practices should be with an actual audience of friends. Practice will help you become familiar with the material, as well as figure out what parts of the presentation work well and what parts need to be rethought. Practice will also make you more comfortable with the presentation.

Finally, practice will help you determine how long the presentation actually lasts, which is important for staying within the scheduled time limit. Note that going through the presentation silently takes much less time than rehearsing out loud, so you should do the latter to determine the actual length of your presentation.

If you are using audiovisual equipment, make sure you know how to operate it before the actual presentation. If at all possible, test out the equipment you will actually be using beforehand. With AV equipment assume the worst case. Plan out how you will complete the presentation if the equipment doesn't work properly. For example, you may wish to bring transparencies of a PowerPoint presentation in case the hardware or software doesn't work. The more sophisticated the presentation, the more things can go wrong. If you plan for the worst, you probably won't need your contingency.

Giving the Presentation

If you have prepared as described in the previous sections, then giving the research presentation will not be too difficult. You may feel nervous, which is normal. Except in extreme cases, the audience will not notice nervousness. In any case, as you get into the presentation, the nervousness will diminish.

Think of the presentation as a conversation about your research, rather than as a speech. These people are here because they are interested in your topic and want to hear what you have to say. You, in turn, want to find out what they think and hear their suggestions for improvement.

You should dress professionally for your presentation. Bring a few copies of your paper, since some people may well request it. If you run out of copies you can always take their names and e-mail addresses and send them the paper later.

It's a good idea to arrive early for the presentation so you won't be rushed setting things up and getting comfortable with the room. Introduce yourself to the other speakers and discussants. If you're presenting with PowerPoint and using someone else's equipment, see if you can load your presentation before the session starts.

As you begin the presentation, stand up so you can project your voice better. Look around the audience and make eye contact with them. Eye contact is important because it engages the audience. It is sometimes helpful to select one person in the left of the audience, one person in the center, and one person on the right to serve as your targets. From time to time in the presentation, remind yourself to speak to each of these persons as you pan the audience with your gaze. If you know your material, this will be no more difficult than riding a bicycle and chewing gum.

Start by introducing yourself and the title of your presentation. As you present your remarks, speak more slowly than you think you have to. Remember, you are intimately familiar with this material, but to your audience it is new. Vary the speed and loudness of your voice; no one likes a monotone speaker.

Think about your body language. Stand in a relaxed manner, rather than stiffly. Don't cross your arms. Don't lean excessively on the podium. Don't fidget with your hands or anything else—it distracts the audience. You don't need to stand in one place throughout the presentation. Feel free to move as much as possible—this will reengage your audience's interest. It will also make you more comfortable and make the occasion seem less formal.

As you speak, avoid filler words like *umm*, *aaah*, and so on. If you have practiced your presentation enough you will know the material so you won't have spaces to fill. In any case, speech experts point out that silence is less distracting to the audience than filler words.

It is unprofessional to exceed the time limit. If your presentation is overly long, at best you will have to speed up, possibly omitting explanations, which will reduce the audience's understanding. At worst, you may be stopped before you conclude.

End the presentation with a simple “thank you,” and then sit down. Let the chair choose to recognize any questions from the audience.

After you complete your presentation, it is likely that you will need to field some questions from the audience. Most will be straightforward, asking for clarification of some point that wasn’t completely understood. Some questions may raise issues you hadn’t considered before. This is a good thing because it will help improve your paper. Take time to think before responding. If you hadn’t considered the issue before say so and thank the person for bringing it up. If, after some thought, you can’t think of a good response, that’s fine. Simply say, you’ll have to think about it further.

The Role of the Discussant

It is common practice at professional conferences for research presentations to be followed by remarks from a discussant. The function of the discussant is twofold: first, to offer a well-thought-out, educated reaction to a research paper, but second, and more importantly, to give the author helpful feedback to improve the paper. In this section, we discuss how to serve effectively as a discussant.

The discussant reads the paper in advance and prepares comments on it.⁵ Ideally, the discussant will be an expert in the field. If you are not, you may need to review the literature in order to be able to talk intelligently about the paper. A good place to start is the key references cited in the paper, especially previous studies that are surveys of the field.

The discussant usually has only a few minutes to comment on a paper, so how should you use that time? To begin, you should briefly summarize the paper and highlight the contribution it attempts to make to the literature. What does the paper try to accomplish? How does it go about doing that? What results are reported? To what extent did the paper succeed in its goal? Often, the discussant can provide a better summary of the paper than the author does, since he or she is seeing it dispassionately and isn’t worried about the details of the work.

What can you suggest to improve the paper? This is where the discussant can really help the author. Could any parts benefit from more detailed explanation? Was there any part that you were unable to understand? Were there any problems with the data that weakened the paper’s argument? Were there any problems in the statistical tests employed? Were there any errors in interpreting the results? Are there any previous studies that the author might benefit from reading?

Specific suggestions are always more helpful than general comments. For example, compare, “I liked this paper” with “A better explanation of where the data came from would make the evidence more convincing.”

The content of the paper is more important than the writing, but presentation matters too to the extent that it detracts from the argument.

In the final analysis, what contribution does the paper make to understanding the issue or problem it seeks to explore? One last point: it is a professional courtesy to give a copy of your comments to the author when you’re done.

SUMMARY

- Researchers disseminate their research in the form of written reports and oral presentations.
- The purpose of research reports and presentations is to present a convincing argument that is supported by logic and empirical evidence.
- A research report includes a title, an introduction and survey of the literature, a theoretical analysis of the issue or problem, empirical testing of the hypothesis, conclusions, and references.
- The theoretical analysis in an empirical research paper is more important than one might think because it is the theory that provides the causal interpretation of the empirical results. Testing without theory can also show correlation at best.
- Interpreting empirical results is more important than presenting them.
- The conclusion should make the case for the research argument: to what extent did the empirical testing confirm the hypothesis?
- A research presentation is best thought of as a summary of a research report. As such, it should be given at a lower technical level and with fewer details.
- The key to minimizing presentation anxiety is to carefully plan the presentation and to practice enough times so that you know the material.
- Never read your presentation!
- To be most effective, you should complement your presentation with visual aids.

NOTES

1. One caveat is thesis writers who have supervisors who may require them to demonstrate mastery of basic concepts.
2. For more detailed discussion of this, see the last section in Chapter 9, “Constructing a Data Appendix to Your Research.”

3. Ethridge (1995) makes the distinction between findings, which are the results of your research, and conclusions, which go beyond your particular project.
4. As noted earlier, the exception to this is in a thesis presentation.
5. Sometimes the author delivers the paper so late that the discussant doesn't have enough time to adequately review it. In such cases, it is perfectly appropriate to indicate that during your comments to the audience.

SUGGESTIONS FOR FURTHER READING

Booth et al. (1995)—How to draft and revise a research paper; not targeted toward economics but useful nonetheless.

Ethridge (1995), especially *Chapter 10*—How to write a research report in economics; designed for graduate students but very readable and useful for undergraduates too.

McCloskey (2000)—Classic though opinionated style guide for writing economics papers.

McCuen et al. (1993), *Chapter 7*—Addressed to scientists and engineers but still relevant to economists.

Thomson (2001)—Addressed to new Ph.D.s but advanced undergraduates may derive some insights about scholarly writing in economics.

Wyrick (1994), *Chapter 12*—How to write an undergraduate research paper in economics.

EXERCISES

1. Create an outline of the argument you will make in your research paper: What is the topic? Why is it worthy of study? In what ways has the literature on the topic failed to completely solve the research problem or answer the research question? How does your research attempt to correct this deficiency? What theory does your research apply? What hypothesis did you obtain from this theory and how did you obtain it? How did you test the hypothesis? What data did you use? To what extent did the test results confirm the hypothesis?
2. Write a survey of the literature on your research topic. Organize your survey so that it provides a justification of the approach you will take in your research.
3. Summarize the theoretical analysis of your research, where the objective is to explain how the theory led to your hypothesis.

4. Write a summary of the empirical tests you performed on your research. What testing methodology did you use? What data did you use? What empirical model did you use? What empirical results would be consistent with the hypothesis? Present and evaluate the actual results. To what extent did they support the hypothesis?
5. Prepare a five-minute presentation on your research paper.
6. Prepare a ten-minute presentation on your research paper.
7. How would the ten-minute presentation differ from the five-minute presentation?

Rubric for Grading Research Papers

1. **Abstract of the Paper**—Does the abstract adequately summarize the author's argument?
2. **Quality of the Research Question**—Does the introduction persuade the reader that the paper investigates a good research question? Is the question interesting and significant?
3. **Quality of the Literature Survey**—Does the literature survey identify the major studies to date on the topic? Does it only review studies that are relevant to the current paper's research question? Does it summarize them succinctly and identify their weaknesses? Does it show how this research relates to the literature, by improving in some way on what has been done previously?
4. **Solid analysis of the issue or problem**—Does the paper display a clear understanding of how to apply economic theory to the issue or problem? Does the analysis lead to the hypothesis of the paper? Is there a clear statement of the research hypothesis?
5. **Valid empirical testing approach**—Is the empirical methodology a reasonable test of the hypothesis? Is the data adequate for the test?
6. **Reasonable interpretation of the results**—Are the results displayed appropriately? Are the results correctly interpreted? Do they support the hypothesis? Could they be interpreted in some way that doesn't support the hypothesis.
7. **Overall Argument**—Based on the evidence provided and interpreted by the author, is the conclusion convincing?
8. **Tables and Graphs**—Are they used effectively?
9. **Reference List**—Is the reference list complete, and does it use the correct citation style?
10. **Data Appendix**—Is the data appendix complete, correct, and adequately cited?

Glossary

abstract – Summary of the argument in a scholarly work. Usually abstracts are limited to 100–200 words.

ad hoc model – A mathematical model that is not derived from optimizing principles. Instead, the hypothesized relationships come from common sense or experience.

annotated bibliography – List of references on a topic, each of which is briefly described and evaluated. An annotated bibliography also includes complete bibliographic information on each reference. Compare with *critical review* and *research abstract*.

annualized growth rate – What the cumulative growth rate over a year would be of an actual growth rate over a shorter period of time. For example, what a monthly growth rate would be if it continued over twelve months.

argument – An assertion or claim supported by reasons or evidence. See also *sound argument*.

auto, or serial, correlation – Common problem of regression analysis. Autocorrelation means that the errors in each observation of

the data sample are dependent on or correlated with each other. Autocorrelation artificially inflates the reported *t* scores, so that explanatory variables may appear to be statistically different from zero when they are not.

Boolean searching – Keyword searching using the Boolean operators AND, OR, or NOT to narrow or widen the search.

browsing – Searching for literature by manually examining documents or bibliographies for references. Contrast with *keyword searching*.

causal validity – The degree to which a test indicates that a relationship is causal, rather than merely correlational. For example, to what extent can one be sure that *X* causes *Y*, rather than the converse, or that both *X* and *Y* are determined by some third variable? Causal validity is one dimension of *internal validity*.

change, or first difference – The difference in values of two adjacent observations. With time-series data, it is the subsequent value less the prior value.

coefficient estimates – See *estimated coefficients*.

confounding, or control, variables – In a hypothesized relationship, confounding variables are other variables that affect the variable you are trying to explain. In order to correctly assess the hypothesized relationship, you need to control for the effects of those other variables. For example, if you believe that consumer spending depends on income, to assess that relationship, you must control for the effects of other variables, such as interest rates or savings, which might also affect consumer spending. See also *multiple regression analysis*.

correlation – A measure of the degree of linear association between two variables. The correlation coefficient ranges from positive one (a perfect positive relationship between the two variables) to zero (no relationship between the two variables) to negative one (a perfect negative relationship between the two variables.) Compare with *covariance*.

covariance – A measure of how two variables vary together. It is related mathematically to the standard deviation and the variance. Compare with *correlation*.

critical review – The summary and evaluation of the argument of a research report, along with complete bibliographic information to identify the report. Compare

with *annotated bibliography* and *research abstract*.

critical thinking – The evaluation of competing arguments based on their logic and evidence.

cross-section data – Different observations of a variable at the same point in time; for example, average personal income across the fifty states in 2004. Contrast with *longitudinal data* and *time-series data*.

deductive reasoning – Reasoning that starts from one or more general principles and derives specific predictions from them. See also *valid deduction*. Contrast with *inductive reasoning*.

dependent variable – The concept we are trying to explain in a hypothesized relationship. Contrast with *explanatory, or independent, variable*.

derived data – Data actually used in empirical testing, which have been manipulated from the source data. Contrast with *raw data*.

descriptive statistics – Statistics that summarize the underlying data. Descriptive statistics include measures of central tendency, such as the *mean*, *median*, and *mode*, and measures of dispersion, such as the *standard deviation* and *variance*.

directories – Hierarchically organized catalogs of information on the Internet on a variety of subjects. See also *virtual libraries*.

dummy variable – A constructed variable in a regression analysis to capture qualitative, usually binary factors such as gender (male/female). Dummy variables have values of one (e.g., male) or zero (e.g., not male). Dummy variables can be used in the construction of an *interaction variable*.

economic, or scientific, significance – The quality of a relationship that is large enough to matter in the real world. See also *size of an estimated coefficient*. Contrast with *statistical significance*.

editing – Correcting the grammar, mechanics, and style of a piece of writing.

endogenous variable – A variable that is determined inside the model being studied. Contrast with *exogenous variable*.

estimated coefficients – The products of *regression analysis*, the estimated coefficients are estimates of the relationship between each *explanatory variable* and the *dependent variable*, *ceteris paribus*. See also *sign of an estimated coefficient*.

explanatory, or independent, variable – A concept that is hypothesized to explain, affect, or cause the *dependent variable*. Indepen-

dent variables can be *exogenous* or *endogenous*. Contrast with *dependent variable*; see also *multiple regression analysis*.

exogenous variable – A variable that is determined outside the model being studied. As such, it is treated as a given in the model. Contrast with *endogenous variable*.

experimental research – Research methodology that is usually characterized by two sample groups: an intervention or treatment group, and a control group. The control group is designed to be in every way identical to the treatment group except for the treatment. Thus, if the experiment between the two groups yields different outcomes, then that difference can be ascribed to the treatment. Contrast with *survey*, or *nonexperimental, research*.

external validity – The degree to which the results of a study can be generalized to other situations, applications, or circumstances. Contrast with *internal validity*.

focus – A feature of good writing, where the point of the document is crystal clear to the reader and the argument does not wander from this point.

good argument – See *sound argument*.

growth rate – See *rate of change*.

heteroskedasticity – A common problem of regression analysis that occurs when the errors in each observation of the data sample, while independent from one another, have different variances. Heteroskedasticity artificially inflates the reported *t* scores, so that explanatory variables may appear to be statistically different from zero when they are not.

implicit price deflator – A price index constructed by dividing the real value of a concept into the nominal value, making it an implicit price index, or deflator. See also *price index* and *unit price*.

independent variable – See *explanatory variable*.

inference – A conclusion reached as the result of reasoning logically about facts and relationships. See also *warranted inference*.

index number – A unit-free weighted average designed to compare trends in prices or quantities since some base period. The *raw data* are transformed so that the value of the base period is 100. Two examples of index numbers include *price indices* and *quantity indices*.

inductive reasoning – Reasoning that infers a general rule from one or more specific cases or situations. Contrast with *deductive reasoning*.

instrument validity – The degree to which a test instrument measures what it purports to. Instrument validity is one dimension of *internal validity*.

interaction term – A constructed variable in a regression analysis designed to capture the effects of two explanatory variables, for example, age and race, interacting. The interaction variable is constructed by multiplying the two explanatory variables together. See also *squared term* and *dummy variable*.

internal validity – The degree to which an impact observed (*Y*) can be attributed to the study variable (*X*); in other words, given the assumptions and evidence, one can deduce that *X* causes *Y*. See also *instrument validity*, *relationship validity*, and *causal validity*. Contrast with *external validity*.

keyword searching – Searching for literature using search engines on the Internet or on specialized electronic databases. Compare with *phrase searching*. Contrast with *browsing*.

knowledge – The meaning or interpretation of evidence.

level of significance – The risk that a researcher is willing to take that the *null hypothesis* will be rejected when it is true. Commonly accepted levels of significance are

1 or 5 percent. A 10 percent level of significance is about the highest most researchers will accept. See also *Type 1 error* and *Type 2 error*.

literature (on a subject) – Studies to date on a subject, both published and in working papers.

literature survey or **review** – The process of identifying the major studies to date on a subject, or a written summary of those studies.

logical fallacy – An argument that is flawed because the conclusion does not actually follow from the reasons stated, even though the argument is phrased in a way that makes you think it does.

longitudinal, or **panel**, **data** – A sample of data that is essentially a cross section followed over time, where strictly speaking the cross section contains the same individuals in each time period. Otherwise, the sample is described as a pooled cross section. Contrast with *cross-section data* and *time-series data*.

maintained hypothesis – The theoretical prediction of a model; what one hopes to prove when conducting a statistical test. If the evidence rejects the *null hypothesis*, then it confirms the maintained hypothesis.

mean – A measure of central tendency, computed as the arith-

metic average. Contrast with *median* and *mode*. See also *descriptive statistics*.

median – A measure of central tendency defined as the middle value of a data sample with half the observations above the median and half the observations below. Contrast with *mean* and *mode*. See also *descriptive statistics*.

mode – A measure of central tendency defined as the most frequent value. Contrast with *mean* and *median*. See also *descriptive statistics*.

moving average – Data-smoothing technique that replaces the actual observation in each period with an average of the *n* – 1 preceding data points.

mathematical reasoning – An approach to theory development based on mathematical manipulation of premises to derive formal hypotheses.

micro data set – Observations of individual economic agents such as individuals, households, or firms.

multicollinearity – A common problem of regression analysis that occurs when two or more explanatory variables are highly correlated. As a result, the reported *t* statistics are artificially low, making it more difficult to prove statistical significance.

- multiple regression analysis** – A statistical method of estimating hypothesized relationships while controlling for the effects of *confounding variables*. Regression is a technique for estimating the independent influences of each explanatory variable, and thus statistically maintaining *ceteris paribus*.
- narrative reasoning** – An approach to theory development based on brainstorming about the topic.
- nominal data** – Values expressed in terms of actual market prices. Nominal is also known as current dollars. Contrast with *real data*.
- nonscience** – Fields of study in which arguments are assessed through logic and normative standards rather than empirical evidence. Contrast with *science*.
- null, or statistical, hypothesis** – The hypothesis that is actually examined in a statistical test—for example, that there is no relationship between two variables. Contrast with *maintained hypothesis*.
- optimizing model** – A mathematical model derived explicitly by assuming that economic agents maximize or minimize an objective function such as profits, utility, or costs.
- organization** – A feature of good writing, where a series of nested arguments lead logically to the thesis as a conclusion.
- parameter estimates** – See *estimated coefficients*.
- per capita** – Per person. Some quantity (e.g., GDP) divided by the population gives the quantity per capita (e.g., GDP per capita).
- percentage** – The *proportion*, or share, expressed as a fraction of one hundred. For example, one-half is 50/100, or 50 percent.
- percent change** – A change expressed as a percentage. See also *rate of change*.
- phrase searching** – Keyword searching on an exact phrase.
- plagiarism** – Taking credit for someone else's words or ideas, even when it's unintentional.
- popular literature** – Summaries of scholarly studies, often written by journalists. Contrast with *scholarly literature*.
- power of a statistical test** – The probability of correctly rejecting a null hypothesis when it is not true.
- prewriting** – Developing a written argument by exploring ideas and relationships. Also known as exploratory writing.
- price index** – The *index number* that is essentially a weighted average of the prices of some type of good or

service, where the weights tend to be either the corresponding quantities or the budget shares in a given year. See also *implicit price deflator*. Contrast with *quantity index*.

primary data source – The person or organization that collected or constructed a data set, often using surveys. See also *secondary data source*.

proportion – The share or fraction of a whole. See also *percentage*.

quantity index – The *index number* that is essentially a weighted average of the quantities of some type of good or service, where the weights are the corresponding prices in a given year. Contrast with *price index*.

random sample of data – A sample for which every item in the population has an equal chance of being selected. An equivalent definition is that every possible sample of the same size n has an equal chance of being the sample selected. A nonrandom sample of data doesn't correctly reflect the underlying population of data; for example, there could be more above- or below-average observations than occur proportionately in the population. See also *sampling error*.

random variation – The variation in some data that is not systematically explainable.

range – A measure of dispersion of a data sample indicated by the lowest value in the sample and the highest value. See also *standard deviation*, *variance*, and *descriptive statistics*.

raw data – The data as they exist at the source. Contrast with *derived data*.

rate of change – A change expressed as a percentage; for example, the change divided by the initial value. See also *percent change*.

real data – Values expressed in terms of prices as they existed in a base period. Real data exclude the effects of inflation since the base period. Real dollars are also known as constant dollars, or 1992 dollars, where 1992 is the chosen base year. Contrast with *nominal data*.

relationship validity – The extent to which one can conclude on the basis of an empirical test that there is a statistical relationship. Relationship validity is one dimension of *internal validity*.

research – The creation of knowledge.

research abstract – A summary of the argument made by a research report, along with complete bibliographic information to identify the report. See also *annotated bibliography* and *critical review*.

research hypothesis – The proposed answer to a research question. The hypothesis is the result deduced from a researcher's analysis. It is what is tested in an empirical research project.

research question – The specific focus of a research project.

research topic – The broad subject area a research project will investigate.

revising – Reviewing the information in an argument, rethinking the way it is organized, and looking for new patterns of meaning; more than just editorial correction.

sampling error – The result of a nonrandom sample of data where the above-average values of some observations are not offset by the below-average values of other observations. As a consequence, the data appear to show a relationship that does not in fact exist. See also *random sample*.

scholarly literature – Studies done by experts in a field and published in professional journals or books. Contrast with *popular literature*.

science – Fields of study in which arguments are empirically testable. Contrast with *nonscience*.

scientific method – The set of procedures for drawing valid, reli-

able, and objective conclusions about scientific hypotheses.

seasonal adjustment – Data-smoothing technique by which normal seasonal fluctuations in the data have been removed. The results show any divergence from the normal seasonal pattern.

secondary data source – An individual or organization that publishes data that it did not produce. Often the data are easier to use than the primary data. Contrast with *primary data source*.

significance (of a research topic or question) – The scientific importance or weight of a research topic or question.

significance level – See *level of significance*.

sign of an estimated coefficient – Important for testing hypotheses about relationships between variables, the (algebraic) sign shows whether a relationship is positive or negative. See also *estimated coefficient* and *size of an estimated coefficient*.

simultaneous equations bias – A common problem of regression analysis. In an ordinary least squares regression, all the *explanatory variables* are assumed to be *exogenous* to the model. In other words, the explanatory variables are hypothesized to affect the dependent variable, but the

converse is not true: the dependent variable should have no effect on the independent variable. If causation runs both ways, the result is simultaneous equations bias and the coefficient estimates will be biased.

size of an estimated coefficient – The magnitude of a hypothesized relationship; the size is important for assessing the *economic significance* of a relationship. See also *estimated coefficient* and *sign of an estimated coefficient*.

Social Science Citation Index – A bibliographic search tool that allows backward and forward searches chronologically. That is, it identifies all the studies cited by a given study as well as all the studies that cite it.

sound argument – A valid deduction in which all the premises are true. See also *argument*.

source data – See *raw data*.

specification error – A common problem of regression analysis, where the empirical model is incorrect, for example, because of omitting relevant variables or including irrelevant ones. In the former case, the estimated coefficients will be biased and the *t* statistics will be artificially reduced; in the latter case, the coefficients are unbiased but the *t* statistics are inflated.

squared term – A constructed variable in a regression analysis used to turn a quadratic relationship that illustrates increasing or diminishing returns into a linear relationship. The variable is constructed by squaring the base factor, for example, $\text{Age_squared} = \text{age} * \text{age}$. See also *interaction term*.

standard deviation – A measure of dispersion of a data sample defined approximately as the average amount by which a data point differs from the mean value of the sample. See also *variance* and *descriptive statistics*.

standard error – A measure of dispersion in a data sample, normalized by the size of the sample. The standard error is defined as the ratio of the *standard deviation* to the square root of the sample size.

statistical hypothesis – See *null hypothesis*.

statistical hypothesis testing – Using methods of statistical inference to test hypotheses about relationships between variables. Typically, statistical methods are used to differentiate between alternatives called the *null hypothesis* and the *maintained hypothesis*. Commonly used tests include the *t*-test and the *F*-test. See also *statistical significance*.

statistical significance – The quality of a statistical relationship that is strong enough to reject the conclusion that it is spurious, due entirely to random chance in the data. See also *statistical hypothesis testing*. Contrast with *economic significance*.

survey, or nonexperimental, research – Survey research involves the passive observation and analysis of events as they occur in nature, while using statistical methods to account for outside factors. Contrast with *experimental research*.

time-series data – Different observations on the same variable at different points across time; for example, U.S. personal income for the period 1954–2004. Contrast with *cross-section data* and *longitudinal data*.

time trend – A constructed variable in regression analysis used to capture the effects of technological improvement or other trends over time.

Type 1 error – Wrongly rejecting the *null hypothesis*. Contrast with *Type 2 error*.

Type 2 error – Wrongly accepting the *null hypothesis*. Contrast with *Type 1 error*.

unit value – An estimate of the price of something defined as the value of shipments divided by the volume of shipments, or the nominal value divided by the real value. Compare with *implicit price deflator*.

valid deduction – A deduction in which the conclusion must follow from the premises.

validity – See *internal* or *external validity*.

variance – A measure of dispersion of a data sample defined as the square of the *standard deviation*. See also *descriptive statistics*.

virtual libraries – Internet directories compiled by librarians or other experts. As such, they tend to have better sources of information than *directories*.

warranted inference – A conclusion that follows from the evidence. See also *inference*.

warrants – Unstated or underlying assumptions on which an argument stands.

writing to learn – The idea that writing is more than a product; it is also a process or a tool for working through and making sense out of ideas you don't fully understand.

Reference List

Ackermann, Ernest, and Karen Hartman. 1998. *Searching & Researching on the Internet and the World Wide Web*. Wilsonville, OR: Franklin, Beedle & Associates.

Bauer, Susan Wise. 2003. *The Well-Educated Mind: A Guide to the Classical Education You Never Had*. New York: W.W. Norton.

Bean, John C. 1996. *Engaging Ideas: The Professor's Guide to Integrating Writing, Critical Thinking, and Active Learning in the Classroom*. San Francisco: Jossey-Bass.

Beck, E. Catesby. 2003. Are consumers more interested in financing incentives or price reductions? *Issues in Political Economy* 12: 55–64.

Best, Joel. 2001. *Damned Lies and Statistics: Untangling Numbers from the Media, Politicians, and Activists*. Berkeley: University of California Press.

Blaug, Mark. 1992. *The Methodology of Economics or How Economists Explain*. Cambridge, UK: Cambridge University Press.

Booth, Wayne C., Gregory C. Colomb, and Joseph M. Williams. 1995. *The Craft of*

Research. Chicago: University of Chicago Press.

Brown, Charles, Greg J. Duncan, and Frank Stafford. 1996. Data watch: The panel study of income dynamics. *Journal of Economic Perspectives* 10 (Spring): 155–168.

Clayton, Gary E., and Martin Gerhard Giesbrecht. 2001. *A Guide to Everyday Economic Statistics*. New York: McGraw-Hill/Irwin.

Cohen, A. J., and J. Spencer. 1993. Using writing across the curriculum in economics: Is taking the plunge worth it? *Journal of Economic Education* 24 (Summer): 219–230.

DeLoach, Stephen B. 2001. More evidence in favor of the Samuelson-Balassa hypothesis. *Review of International Economics* 9 (May): 336–342.

Ehrenberg, Ronald, John Cheslock, and Julia Epifantseva. 2001. Paying our presidents: what do the trustees value? *Review of Higher Education* 25 (Fall): 15–37.

Epstein, Richard L., and Carolyn Kernberger. 2005. *The Guide to Critical Thinking in Economics*. Mason, OH: Thomson South-Western.

- Ethridge, Don. 1995. *Research Methodology in Applied Economics*. Ames: Iowa State University Press.
- Fallen, Melissa. 2000. Meaningful money in a new classical model. Paper presented at the annual meeting of the Eastern Economic Association, Washington, DC.
- Friedman, Milton. 1968. Why Economists Disagree. In *Dollars and Deficits*. Englewood Cliffs, NJ: Prentice-Hall, 1–16.
- Graff, Gerald. 2003. *Clueless in Academe: How Schooling Obscures the Life of the Mind*. New Haven, CT: Yale University Press.
- Gujarati, Damodar N. 2003. *Basic Econometrics*. New York: McGraw-Hill.
- Harvey, Michael. 2000. *The Nuts and Bolts of College Writing*. Indianapolis, IN: Hackett <http://www.nutsandboltsguide.com> (accessed 19 June 2003).
- Hausman, Daniel M. 1989. Economic methodology in a nutshell. *Journal of Economic Perspectives* 3 (Spring): 115–128.
- Huff, Darrell. 1954. *How to Lie with Statistics*. New York: W.W. Norton.
- Kennedy, Peter. 2003. *A Guide to Econometrics*. Cambridge: MIT Press.
- Knoblauch, C. H., and L. Brannon. 1983. Writing as learning through the curriculum. *College English* 45 (September): 465–474.
- Krugman, Paul. 1979. Increasing returns, imperfect competition, and international trade. *Journal of International Economics* 9 (November): 469–479.
- Lanham, Richard A. 1992. *Revising Prose*. New York: MacMillan.
- Locke, Lawrence F., Stephen J. Silverman, and Waneen Wyrick Spirduso. 1998. *Reading and Understanding Research*. Thousand Oaks, CA: Sage.
- Machlup, Fritz. 1965. Why economists disagree. *Proceedings of the American Philosophical Society* 109 (February): 1–7.
- Maier, Mark H. 1999. *The Data Game: Controversies in Social Science Statistics*. Armonk, NY: M.E. Sharpe.
- McCloskey, D. 2000. *Economical Writing*. Prospect Heights, IL: Waveland Press.
- McCloskey, D. 1999. Other things equal: Economical writing: An executive summary. *Eastern Economic Review* 25 (Spring): 239–242.
- McCloskey, D. 1998. *The Rhetoric of Economics*. Madison: University of Wisconsin Press.
- McCuen, Richard H., Peggy A. Johnson, and Cynthia Davis. 1993. *Dynamic Communication for Engineers*. New York: American Society of Civil Engineers.
- McLagan, Donald L. 1973. A non-econometrician's guide to econometrics. Federal Reserve Bank of Philadelphia *Business Economics* (May): 38–45.
- Miller, Merton H. and Charles W. Upton. 1974. *Macroeconomics: A Neoclassical Introduction*. Chicago: University of Chicago Press.
- Missimer, C. A. 1995. *Good Arguments: An Introduction to Critical Thinking*. Englewood Cliffs, NJ: Prentice Hall.
- Morgenstern, Oskar. 1979. *National Income Statistics: A Critique of Macroeconomic Aggregation*. San Francisco: Cato Institute. Reprinted from *On the Accuracy of Economic Observations*, 1963, Princeton, NJ: Princeton University Press, Chap. 14.
- Mursell, James. 1951. *Using Your Mind Effectively*. New York: McGraw-Hill.
- Nelson, C. 1989. Skewered on the Eunuch's Horn: The Illusion of Tragic Tradeoff Between Content and Critical Thinking in the Teaching of Science. In *Enhancing Critical Thinking in the Sciences*, ed. L. Crow. Washington, DC: Society of College Science Teachers, 17–27.
- Newman, Sandra, and Joseph Harkness. 2000. Assisted housing and the educational attainment of children. *Journal of Housing Economics* 9 (March–June): 40–63.
- Palmini, Dennis J. 1994. Using rhetorical cases to teach writing skills and enhance economic learning. Paper presented at the Annual Meeting of the Eastern Economics Association, New York, March 18–20.
- Pergamit, Michael R., Charles R. Pierret, Donna S. Rothstein, and Jonathan R. Veum. 2001. Data watch: The national longitudinal surveys. *Journal of Economic Perspectives* 5 (Spring): 239–254.
- Perry, W. 1970. *Forms of Intellectual and Ethical Development in the College Years: A Scheme*. New York: Holt, Rinehart & Winston.
- Petr, Jerry L. 1998. Student writing as a guide to student thinking. In *Teaching Undergraduate Economics: A Handbook for Instructors*, ed. William L. Walstad and Phillip Saunders. Boston: McGraw-Hill/Irwin, 227–243.
- Phillips, John L. Jr. 1996. *How to Think About Statistics*. New York: W.H. Freeman.
- Popkin, Joel. 2000. Data watch: The U.S. national income and product accounts. *Journal of Economic Perspectives* 14 (Spring): 215–224.
- Ramanathan, Ramu. 1995. *Introductory Econometrics: With Applications*. Fort Worth, TX: Dryden.
- Remenyi, Dan, Brian Williams, Arthur Money, and Ethne Swartz. 1998. *Doing Research in Business*

- and Management: An Introduction to Process and Method. London: Sage.
- Romer, David. 1993. Do students go to class? Should they? *Journal of Economic Perspectives* 7 (Summer): 167–174.
- Ruggiero, Vincent Ryan. 1998. *The Art of Thinking: A Guide to Critical and Creative Thought*. New York: Addison Wesley Longman, Chap. 4.
- Samuelson, Robert J. 2002. A war we can afford. *Washington Post*, September 18, A29.
- Schroeder, David A., David E. Johnson, and Thomas D. Jensen. 1985. Reading research reports: A brief introduction. In *Contemporary Readings in Social Psychology*, ed. David A. Schroeder, David E. Johnson, and Thomas D. Jensen. Chicago: Nelson-Hall, 36–42.
- Siegfried, John, Robin Bartlett, W. Lee Hansen, Alan Kelley, Donald McCloskey, and Thomas Tietenberg. 1991. The status and prospects of the economics major, *Journal of Economic Education* 22 (Summer): 197–224.
- Sorokina, Olga. 2003. Executive compensation: The case of liberal arts college presidents. *Issues in Political Economy* 12: 8–24.
- Spector, Paul E. 1981. *Quantitative Applications in the Social Sciences*. Research Design Series, no. 23. Newbury Park, CA: Sage.
- Studenmund, A. H. 2001. *Using Econometrics: A Practical Guide*. New York: Addison-Wesley-Longman.
- Teplin, Albert M. 2001. The U.S. flow of funds accounts and their uses. *Federal Reserve Bulletin* (July), 431–441.
- Thomson, William. 2001. *A Guide for the Young Economist: Writing and Speaking Effectively about Economics*. Cambridge, MA: MIT Press.
- Tobin, James. 1978. A proposal for monetary reform. *Eastern Economic Review* 4 (July/October): 153–159.
- Trelogan, T. K. 2001. Arguments and their evaluation. <http://www.unco.edu/philosophy/arg.html> (accessed June 22, 2004).
- Turabian, Kate L. 1996. *A Manual for Writers of Term Papers, Theses and Dissertations*. Chicago: University of Chicago Press.
- Williams, Willard F. 1984. Unpublished class notes for research methodology in economics, Texas Tech University.
- Wyrick, Thomas L. 1994. *The Economist's Handbook: A Research and Writing Guide*. St. Paul, MN: West Publishing Company.
- Zinnser, W. 1988. *Writing to Learn*. New York: Harper & Row.

Index

- Abstracts, 118, 119 (table)
of research reports, 247
- Accuracy, for persuasive arguments, 59–60
- Ackermann, Ernest, 33, 38, 44
- Acknowledgements, in research reports, 247
- Active voice, for economic writing, 85–86
- Ad hoc models, 131–132, 134
- Ad hominem fallacy, 68
- Advisors, choosing, 24
- AEA (American Economic Association), 33 (table), 33–34
web portal hosted by, 41
- AER (*American Economic Review*), 30, 41
- Affirming the consequent, 67
- AHS (American Housing Survey), 144, 148–149
- Alternative hypothesis, 192, 193, 194
- American Economic Association (AEA), 33 (table), 33–34
web portal hosted by, 41
- American Economic Review* (AER), 30, 41
- American Enterprise Institute, 32
- American Housing Survey (AHS), 144, 148–149
- American Psychological Association (APA) citation style, 46, 47
- AND operator, 37, 39–40
- Annotated bibliographies, 118–119
- Annualized growth rate, 167–168
- Annual Survey of Manufactures (ASM), 144, 148
- APA citation style, 46, 47
- Appeals
to authority, 68
to the people (to the many), 68
to pity, 69
- Appendices, in research reports, 175–176, 180–181, 248, 250
- Arguments, 2
constructing, 39, 63. *See also* Writing as process
definition of, 64
disproving versus proving, 5
evaluating, 3–5, 95–97, 113 (table)
examining, 54–57
flawed, 67–69
identifying, 99 (table)
incorrect, 64
inductive, 58
logical or theoretical reasons for, 111
persuasive, 59–64, 116
reasoning and, 57–59
sound, 57
structure of, 51–54, 53 (fig.)
- Asian Development Bank, 145
- ASM (Annual Survey of Manufactures), 144, 148
- Assertions, proposed by author, 110
- Assumptions
reasonableness of, 115
in reasoning, 112
- Author-date system, 46
- Authority
appeals to, 68
for persuasive arguments, 60
- Autocorrelation, 217–219, 220

Backward linking, 41
 Bartlett, Robin, 25
 Bauer, Susan Wise, 120
 BEA (Bureau of Economic Analysis), 144
 Bean, John C., 1, 7, 15, 24, 50, 72, 96–97, 120
 Begging the question, 67
 Best, Joel, 140, 141, 155, 200, 201
 Bibliographic databases, 35
 Bibliographies, annotated, 118–119
 Blaug, Mark, 7, 9, 24
 BLS (Bureau of Labor Statistics), 145
 Books
 online, citing, 47
 sources of, 42
 Boolean searching, 37, 39–40
 Booth, Wayne C., 1, 3, 6, 7, 15, 16, 18, 24, 25, 61, 64, 65, 73, 77, 78, 79, 80, 85, 90, 236, 238, 260
 Brannon, L., 64
 Brookings Institution, 32
 Brown, Charles, 155
 Browsing, as search strategy, 33 (table), 33–35
 Bureau of Economic Analysis (BEA), 144
 Bureau of Labor Statistics (BLS), 145
 Bush, George W., 68
Business Week magazine, 31
 Capital Account, 146
 Cato Institute, 32
 Causal empiricism, 187–189, 188 (fig.)
 Causal validity, 185, 198–199
 Census Bureau, 144
 CES (Consumer Expenditure Survey), 145, 149
 Chain-type indexes, 177
 Charts, in research reports, 235
 Cheslock, John, 240
Chicago Manual of Style, 46, 118
 Chicago Style, 46
 CIR (Current Industrial Reports), 144, 148
 Citation styles, 46–47
 Clarity
 of thesis, 80
 writing style and, 83–85
 Clayton, Gary E., 155
 Cobb-Douglas production function, 228
 Coefficients
 estimated, see Estimated parameters (coefficients)
 signs of, 209
 Cohen, A. J., 3, 118
 Colomb, Gregory C., 1, 3, 6, 7, 15, 16, 18, 24, 25, 61, 64, 65, 73, 77, 78, 79, 80, 85, 90, 236, 238, 260
 Communicating research results, see Oral presentations; Research reports
 Completeness, for persuasive arguments, 61, 62–63
 Composition (writing), 50–51. *See also* Arguments; Writing as process
 Composition, fallacy of, 68–69
 Conclusions
 drawing, 22
 following from evidence, 53–54, 116
 in research report, 245–246
 Conference Board, 151
 Conferences, undergraduate presentations at, 23, 232
 Conflicts of interest, of authors, 112
 Confounding variables, 196–198, 197 (fig.)
 Congressional Budget Office, 32
Consumer Confidence Index, 151

Consumer Expenditure Survey (CES), 145, 149
 Consumer price index (CPI), 170–171, 171 (fig.), 177
 Convenience sampling, 200
 Correlation, 189
County & City Data Book, 144
 Covariance, 189
 CPI (consumer price index), 170–171, 171 (fig.), 177
 CPS (Current Employment Statistics), 145, 147
 CPS (Current Population Survey), 145, 147
 Critical reading, 95–97
 distinctions for, 97
 of published research, 95–97
 questions to guide, 97, 99 (table), 99–112, 113 (table), 113–117
 of theoretical articles, 122–123
 Critical reviews, 118, 119 (table)
 Critical thinking, 4
 Cross-section data, 142
 Current Account, 146
 Current Employment Statistics (CPS), 145, 147
 Current Industrial Reports (CIR), 144, 148
 Current Population Survey (CPS), 145, 147
 Data
 adequacy of, 114–115
 derived, 175–176
 nominal, 169–170
 raw, 175–176
 real, 170
 seasonally adjusted, 175
 transformations of, for regression analysis, 228
 See also Economic data
 Data appendices, 175–176, 180–181, 250
 Databases, keyword searches and, 35–38, 39–40
 Data manipulation, 164 (fig.), 164–176
 change in variable and, 165, 167–169, 168 (fig.), 169 (fig.)
 data appendixes and, 175–176, 180–181
 data smoothing and, 174–175
 inflation distortion of nominal values and, 172
 level of variable and, 164 (fig.), 164–165
 price indices versus implicit price deflators and, 172
 quantity indices versus real quantities and, 171–172
 real versus nominal magnitudes and, 169–170
 rebasings data series and, 172–174, 174 (fig.)
 Data series
 constructing, steps in, 140–141
 rebasings, 172–174, 174 (fig.)
 Data sets, 159–176
 characteristics of, 142–143, 143 (fig.)
 data manipulation and (*see* Data manipulation)
 developing search strategies for data finding and, 160 (fig.), 160–163
 locating, 21
 macro, 154
 micro, 142, 154
 Data smoothing, 174–175
 Davis, Cynthia, 83, 90, 201, 252–253, 260
 Deductive reasoning, 9, 57–58
 Dependent variable, 204
 qualitative (limited), 230
 Derived data, 175–176
 Descriptive statistics, 187–188

- Development, for good economic writing, 75–76, 76 (fig.)
- DIALOG database, 35
- Directories, Internet, 31
- Disciplines, science versus nonscience, 11–12
- Discussants, 258–259
- Division, fallacy of, 68–69
- Dow Jones Interactive database, 35
- Drafts, revising, 79–82
- Dualistic thinking, 3–4
- Dummy variables, 229–230
- Duncan, Greg J., 155
- Durbin-Watson statistic (DW), 218–219
- Eastern Economic Association, 23, 232
- Eastern Economic Journal*, 113
- EconLit*, 35
- Economagic, 151–152
- Economic Census, 144, 148
- Economic data, 139–154
creation of, 140–141
organizations collecting and publishing, 144–145
primary, major collections of, 145–151
secondary, major collections of, 151–154
sources of, overview of, 157–158
structure of, 141–143
- Economic methodologies, 9
- Economic Policy Institute, 32
- Economic Report of the President (ERP), 151
- Economic Review*, 32
- Economic significance, statistical
significance versus, 212–213
- Economics working papers, 42
- Economic writing, 71–90
economic writing defined and, 71–72
first draft and, 72–79
giving credit for intellectual property and, 77–79
good, features of, 73–76, 74–76 (fig.)
mechanics and, 87–89
revising and, 79–82
steps in writing and, 72, 73
writing style and, 82–87
See also Research reports
- Economist*, 30
- Editing, 87–89
- Ehrenberg, Ronald, 240
- Electronic books, 42
- Empirical testing, 185–186, 199
in research report, 241–245
- Encyclopedias, 31
- Endogenous variables, 221
- Epifantseva, Julia, 240
- Epstein, Richard L., 65
- ERP (Economic Report of the President), 151
- Errors
sampling, 191
specification, 218, 221–222
standard, 195
type 1 and type 2, 194
- Estimated parameters (coefficients), 20
assessing statistical significance of, 211
conformity with theoretical predictions, 209 (fig.), 209–217
plausibility of, 213–216, 215
size of, 211, 214
- Ethridge, Don, 3, 6, 7, 13, 22, 24, 25, 26, 29, 30, 129, 135, 137, 198, 260
- EuroStat, 145, 153
- Evidence
empirical, provided by author, 111–112
for persuasive arguments, 59–63
supporting points with, 83
- Expanded Academic ASAP*, 41
- Experimental methods, 184
- Explanation
of assumptions, clarity of, 115
for persuasive arguments, 60–61
- Explanatory variables, 204
- External validity, 185
- Fallacies, logical, 54, 67–69
- Fallacy of composition (of division), 68–69
- False analogy, 69
- Federal Reserve, 32, 145
- Figures, in research reports, 235
- First draft, 72–79
- First person, in economic writing, 86
- Fisher, Irving, 132
- Flow of Funds Accounts, 145, 146
- Focus, for good economic writing, 73–74, 75 (fig.)
- Forward linking, 41
- FRED II, 145, 152
- Friedman, Milton, 7, 8, 57
- F statistic, 216
- Full-text databases, 35
- GDP, per capita, 164–165
- Giesbrecht, Martin Gerhard, 155
- Goffe, Bill, 41
- Government documents, on Internet, 41–42
- Graff, Gerald, 65
- Graphics, in research reports, 235
- Greenspan, Alan, 67
- Gross domestic product (GDP), per capita, 164–165
- Growth rate, 167
annualized, 167–168
- Gujarati, Damodar, 227
- Hansen, W. Lee, 25
- Harnock, Andrew, 47
- Hartman, Karen, 33, 38, 44
- Harvey, Michael, 85, 90
- Hausman, Daniel M., 7, 9
- Heteroskedasticity, 219–221
- Huff, Darrell, 155, 201
- Hypotheses
adequacy of methodology for testing, 115
alternative (maintained), 192, 193, 194
attempting to refute, 9
definition of, 14
failed, 246
good, characteristics of, 134–136
null (statistical), 192–193, 194, 200
stating, 204–206, 205 (fig.)
- Hypothesis testing
in regression analysis, 206–208, 207 (fig.), 208 (fig.)
statistical, 184–186
- IFS Yearbook*, 152–153
- Implicit price deflators, 172
- Independent variable, 204
qualitative, estimation using, 229–230
- Index numbers, 170–171, 171 (fig.)
- Index of Consumer Sentiment*, 150–151
- Inductive reasoning, 9, 58
- Inference, 53–54
- Inflation, distortion of nominal values by, 172
- Instrument validity, 184–185
- Integrated Public Use Microdata Series (IPUMS), 147

- Interaction terms, 230
- Inter-American Development Bank, 145
- Interlibrary loan services, 42
- Internal validity, 184
- International Financial Statistics (IFS), 152–153
- International Monetary Fund (IMF), 145, 152
- Internet Public Library, 31
- Internet resources, 31
- Inter-university Consortium for Political and Social Research (ICPSR), 152
- Introduction, in research report, 234–236
- Issues in Political Economy*, 23, 232
- Jensen, Thomas D., 93, 121
- Johnson, David E., 93, 121
- Johnson, Peggy A., 83, 90, 201, 252–253, 260
- Joint-BIS-IMF-OECD-World Bank Statistics on External Debt database, 153
- Journal of Economic Literature* (JEL), 34, 41, 94
- Journal of Economic Perspectives* (JEP), 34, 35, 41
- Journals, 23
 - lists of, 41
 - online, citing, 47
 - publishing undergraduate research, 23, 232
 - refereed, 112–113
 - sources of, 41
- JSTOR, 41, 41 (table)
- Kelly, Alan, 25
- Kennedy, Peter, 213
- Kernberger, Carolyn, 65
- Keynes, John Maynard, 3, 69
- Keyword searches, 35–38, 39–40
- Kleppinger, Eugene, 47
- Knoblauch, C. H., 64
- Knowledge
 - creation of, 2–3, 50
 - validity of, 3–5
 - ways of gaining, 24
- Krugman, Paul, 94, 122
- Lanham, Richard A., 84
- Laspeyres indexes, 177
- Level of significance, 193–194
- Level of variables, 164 (fig.), 164–165
- Libraries, 42
 - virtual, 31
- Limited dependent variables, 230
- Linear regression, nonlinear regression versus, 226
- Literature, 19
 - review of, writing, 236–239
- Literature surveys, 19–20, 29–43
 - obtaining resources for, 40–42, 41 (table)
 - of popular versus scholarly literature, 30–32, 32 (table)
 - purpose of, 29–30
 - search strategy for, 32–40
- Locke, Lawrence F., 98, 99, 116–117, 121
- Logical fallacies, 54, 67–69
- Longitudinal data, 142
- Machlup, Fritz, 7, 8
- Macro data sets, 154
- Maier, Mark H., 145, 146, 147, 155
- Main Economic Indicators, of OECD, 153–154
- Maintained hypothesis, 192, 193, 194
- Manual for Writers of Term Papers, Theses and Dissertations*, A (Turabian), 46

- Marshall, Alfred, 3
- Mathematical reasoning, 129–132
 - ad hoc models and, 131–132, 134
 - optimizing models and, 129–130
- McCloskey, D., 9, 25, 52, 64, 65, 73–74, 76–77, 78, 84, 89, 90, 97, 212–213, 260
- McCuen, Richard H., 83, 90, 201, 252–253, 260
- McLagan, Donald L., 227
- Mean, 187–188
- Mechanics, in economic writing, 87–89
- Median, 188, 200
- MetaCrawler, 31
- Meyers, Peter B., 98
- Micro data sets, 142, 154
- Miller, Merton H., 137
- Missimer, C. A., 4, 6, 24, 52, 53, 58, 65, 251
- MLA citation style, 46–47
- Mode, 188
- Modern Language Association (MLA) style, 46–47
- Modigliani, Franco, 132
- Money, Arthur, 7, 128, 137
- Monthly Labor Review*, 145
- Morgenstern, Oskar, 155
- Moving averages, 174
- Multicollinearity, 222–223
- Multiple (multivariate) regression analysis, 198, 206
- Multiplicity, 4–5
- Mursell, James, 96
- Narrative reasoning, 128–129
- National Accounts, of OECD, 153–154
- National Bureau of Economic Research, 42
- National Conference on Undergraduate Research, 23, 232
- National Income and Product Accounts (NIPA), 144, 146
- National Longitudinal Surveys (NLS), 145, 149–150
- Nelson, C., 6, 7
- Nested Boolean logic, 37
- NetEc, 42
- NetLibrary, 42
- News magazines, 30, 31
- Newspapers, 31
- Newsweek* magazine, 31
- New York Times*, 31
- Nominal data, 169–170
- Nominalization, 86
- Nominal values, distortion by inflation, 172
- Nonexperimental methods, 184
- Nonlinear regression, linear regression versus, 226
- Nonscience disciplines, 11–12
- Normal distribution, 188–189
- North American Industry Classification System (NAICS), 147, 148
- Note taking, 117–118
- NOT operator, 37
- Novice researchers
 - autocorrelation and, 220
 - critical *t* statistic and, 210
 - Durbin-Watson statistic and, 219, 220
 - empirical testing and, 199, 246
 - expertise of, 251
 - international comparisons and, 166–167
 - literature searches by, 36
 - measurement units and, 165
 - multicollinearity and, 223
 - oral presentations by, 23, 232, 251, 254, 256
 - plausibility of estimated coefficients and, 215
 - publication by, 23, 232

- reading of research by, 97, 98
 regression problems and, 223
 report writing by, 52, 73, 78, 86, 232–250
 requirements for good research projects and, 13
 research design by, 135
 research questions and, 17–18
 sample regression result interpretation and, 224
 size of estimated coefficients and, 214
 statistical significance and, 211
 theorizing and, 127
 transforming nominal into real variables and, 173
 type 1 and type 2 error and, 194
 value of research conducted by, 16
 Null hypothesis, 192–193, 194, 200
- OCLC (On-Line Computer Library Center), 42
 OCLC FirstSearch database, 35
 OECD (Organization for Economic Cooperation and Development), 145, 153–154
Online: A Reference Guide to Using Internet Sources (Harnock and Kleppinger), 47
 Online citation styles, 47
 On-line Computer Library Center (OCLC), 42
 On-line Writing Lab (Purdue University), 47
 Optimizing models, 129–130
Opus 1, 232
 Oral presentations
 differences from written presentations, 251–252
 discussants and, 258–259
 emphasizing important points in, 256
 giving, 256–258
 practicing for, 256
 preparing, 252–253
 by undergraduates, 23, 232, 251, 254, 256
 visual aids for, 253–255
 Organization
 for good economic writing, 74–75, 76 (fig.), 80–82
 Organization for Economic Cooperation and Development (OECD), 145, 153–154
 OR operator, 37, 39–40
 Outlining
 for economic writing, 76, 83 (fig.)
 using word processing software, 52
- Paasche indexes, 177
 Palmmini, Dennis J., 71
 Panels, 154
 Panel Survey of Income Dynamics (PSID), 150
 Parameters, estimated, *see* Estimated parameters (coefficients)
 Penn World Tables, 153
 Per capita form of variables, 164–165
 Percentage change, 167
 Percentages, 169
 Pergamit, Michael R., 155
 Perry, W., 6, 7
 Persuasive arguments, 59–64
 Petr, Jerry L., 72
 Phillips curve, 187, 188 (fig.)
 Phillips, John L. Jr., 201
 Phrases, as sentence fragments, 88
 Phrase searching, 38
 Pity, appeals to, 69
 Plagiarism, 77–79
 avoiding, 78–79
 types of, 77–78
 Popkin, Joel, 155
- Popular publications, 30–31
 Post hoc, ergo propter hoc fallacy, 68
 Poterba, James, 35
 Power, of tests, 186
 PowerPoint presentations, 253–255
 PPP Equivalent, 166, 177
 Practicing, for oral presentations, 256
 Precision
 for economic writing, 86–87
 for persuasive arguments, 60
 Previous research, reading research reports and, 110–111
 Price indices, 170
 implicit price deflators versus, 172
 Professional conferences, undergraduate presentations at, 23, 232
 Professional publications, 30, 31–32, 32 (table). *See also* Journals
 Proportion, 169
 PSID (Panel Survey of Income Dynamics), 150
 Publications, scholarly (professional; scientific), 30, 31–32, 32 (table). *See also* Journals
 Purchasing power parity equivalent (PPP Equivalent), 166, 177
p values, 196, 210–211
- Qualitative variables
 dependent, 230
 estimation using, 229–230
 Quantity indexes, 170
 real quantities versus, 171–172, 177
 Quotations, caution about use of, 78
- Ramanathan, Ramu, 25
 Random data points, 191
 Random sampling, 160
 Random variation, in human behavior, 189–191, 190 (fig.), 192 (fig.)
 Range, 188
- Rate of change, 167
 Raw data, 175–176
 Reading published research, 93–120
 critical reading for, 95–97
 evaluating research and, 112–113
 format and, 94–95
 note taking and writing abstracts and reviews and, 117–120, 119 (table)
 questions to guide, 97, 99 (table)
 99–112, 113 (table), 113–117
 Real data, 170
 Real quantities, quantity indexes versus, 177
 Reasoning
 author's assumptions in, 112
 deductive, 57–58
 inductive, 58
 mathematical, 129–132
 narrative, 128–129
 warrant-based, 58–59
 Reference lists, in research reports, 248
 Regression analysis, 203–225
 coefficient estimates and, 209 (fig.), 209–216
 data transformations for, 228
 evaluation of test results and, 224
 goodness of fit and, 216–217
 interpreting tests results and, 208–217
 problems common in, 217–223
 stating hypotheses and, 204–206, 205 (fig.)
 statistical versus economic significance and, 212–213
 steps in, 203, 204 (fig.)
 testing hypotheses and, 206–208, 207 (fig.), 208 (fig.)
 Remenyi, Dan, 7, 128, 137
 Reporting, 23
 Representativeness
 for persuasive arguments, 61, 62
 of samples, 160

- Research advisors, choosing, 24
- Research design, 135, 183–199
 - causal validity and, 198–199
 - confounding variables and, 196–198, 197 (fig.)
 - data analysis and, 186–189
 - key issues of, 183–186
 - parts of, 21–22
 - random variation in behavior and, 189–191, 190 (fig.), 192 (fig.)
 - simple statistical hypothesis testing and, 194–196
 - statistical methods and, 192–194, 193 (fig.)
- Research hypotheses. *See* Hypotheses
- Research papers
 - types of, 94–95
 - See also* Research reports
- Research problems
 - identification by scholars, 24
 - theoretical analysis of, 20–22
- Research process, 11–23
 - for economics, steps of, 12–23, 13 (table)
 - in science versus nonscience disciplines, 11–12
- Research proposals
 - components of, 26–27
 - questions for evaluating, 28
 - writing, 26–27
- Research questions
 - definition of, 14
 - developing, 14–16, 15 (table), 18–19
 - economic, 18
 - examples of, 17–18, 19
 - identifying, 110
 - significance of, 16, 18
- Research reports, 23, 231–250
 - abstract of, 247
 - acknowledgements in, 247
 - appendices in, 175–176, 180–181, 248, 250
 - components of, 233, 233 (fig.)
 - conclusions in, 245–246
 - empirical testing of analysis and, 241–245
 - graphics in, 235
 - introduction to, 234–236
 - literature review in, 236–239
 - reading (*see* Reading published research)
 - reference lists in, 248
 - style for, 248–249
 - table of contents of, 247
 - theoretical analysis in, 239–241
 - title of, 247
- Research topics
 - applying theory to, 126
 - definition of, 14
- Research university libraries, 42
- Resources for Economists (RFE)
 - website, 41, 161
- Results
 - interpreting, 22
 - reporting, 23
- Revising, 79–82
- RFE (Resources for Economists)
 - website, 41, 161
- Romer, David, 99–112, 114, 115, 116
- R² statistic, 216, 217
- Ruggiero, Vincent Ryan, 97, 115, 116, 121
- Run-on sentences, 89
- Sample data, 141
- Sample size, 160
- Sampling error, 191
- Samuelson, Robert J., 54–56, 57
- SCF (Survey of Consumer Finances), 145, 151
- Scholarly publications, 30, 31–32, 32 (table). *See also* Journals
- Schroeder, David A., 93, 121
- Science disciplines, 11–12

- Scientific method, 12
 - for economics, steps in, 12–23, 13 (table)
- Scientific publications, 30, 31–32, 32 (table). *See also* Journals
- Search engines, 31
 - keyword searches and, 35–38, 39–40
- Search strategies, 32–40
 - basic, 38 (table), 38–39
 - browsing as, 33 (table), 33–35
 - for data finding, 160 (fig.), 160–163
 - keyword searching as, 35–38, 39–40
- Seasonally adjusted data, 175
- Sentence fragments, 88
- Sentences
 - complete, 88
 - run-on, 89
 - structure of, 83–84
- Shares, 169
- SIC (Standard Industrial Classification) system, 148
- Siegfried, John, 25
- Significance
 - economic, statistical significance versus, 212–213
 - of research questions, 16, 18
 - statistical (*see* Statistical significance)
- Signs, of coefficients, 209
- Silverman, Stephen J., 98, 99, 116–117, 121
- Simultaneous equations bias, 221
- SOAS Economic Digest, 232
- Social Science Citation Index (SSCI), 35–37
- Social Science Research Network, 42
- Sorokina, Olga, 240
- Sound arguments, 57
- Special pleading fallacy, 67
- Specification error, 218, 221–222
- Spector, Paul E., 12
- Speculation, 24
- Spell-checking software, 89
- Spencer, J., 3, 118
- Spirduso, Waneen Wyrick, 98, 99, 116–117, 121
- Squared terms, 228
- SSCI (*Social Science Citation Index*), 35–37
- Stafford, Frank, 155
- Standard deviation, 188
- Standard error, 195
- Standard Industrial Classification (SIC) system, 148
- State & Metropolitan Area Data Book*, 144
- Statistical Abstract of the United States*, 144
- Statistical hypotheses, 192–193, 194, 200
- Statistical hypothesis testing, 184–186
- Statistical method, selecting, 22
- Statistical significance
 - economic significance versus, 212–213
 - of estimated coefficients, novice researchers, 211
 - importance of, 216–217
 - level of, 193–194
- Statistics, descriptive, 187–188
- STAT-USA, 152
- Straw man fallacy, 67
- Studenmund, A. H., 227
- Student Economic Review, The*, 232
- Survey methods, 184
- Survey of Consumer Finances (SCF), 145, 151
- Survey of Current Business*, 144, 145
- Survey Research Center (University of Michigan), 150–151
- Swartz, Ethne, 7, 128, 137
- Tables, in research reports, 235
- Tables of contents, of research reports, 247
- Teplin, Albert M., 155